UNIVERSITY OF PENNSYLVANIA

THE WHARTON SCHOOL

INSR 205

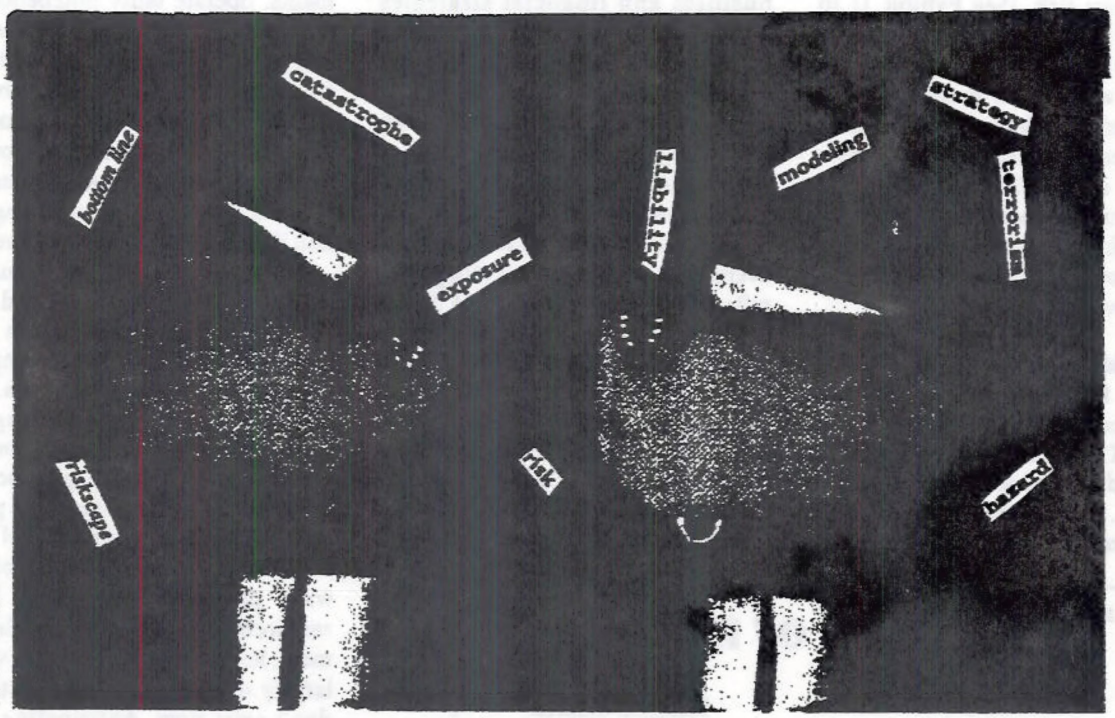
Professor Nyce | Spring 1998

---

Table of Contents

# Toward a Universal Language of Risk

by Jerry A. Miccolis




JAMES YANO

ORGANIZATIONS AROUND THE WORLD ARE FUNDAMENTALLY CHANGING the way they view risk. Senior managers are looking at the entire range of their exposures and are beginning to examine the way independent risks interact and the individual and combined effects they have on the organization's bottom line. Unfortunately, the language used to describe risks is becoming muddied. Driven in part by the increased attention being given to pure financial risk in most organizations, the definitions of *risk* and *risk management* are changing. What a banker calls risk management is likely to be far different from what a traditional risk manager defines. As businesses expand the definition of risk to include both "hazard" risks (such as property and casualty risks) and "nonhazard" risks (such as financial and new product risks), there should

*Jerry A. Miccolis, FCAS, MAAA, leads Tillinghast–Towers Perrin's worldwide risk management practice. He is based in the firm's Parsippany, NJ, office.*

**In order for senior managers to have a complete grasp of all-encompassing risk as it affects their businesses, they need to communicate the varieties of risk in a common language.**

be a convergence of the treasurer's and the risk manager's definition of risk. This more complete definition is closer to the way a typical CFO or CEO would view risk. [See Exhibit 1] In order for senior managers to have a complete grasp of all-encompassing risk as it affects their businesses, they need to communicate the varieties of risk in a common language. Only then can they approach risk holistically, with an understanding of how the risks work independently and together, and how they could affect the bottom line when combined.

### The Graphic Model

To lay a foundation for discussion among senior managers across disciplines, Tillinghast-Towers Perrin developed a graphic way of illustrating risk. This model has several advantages: it can represent risks of all types; it can encompass any number of variables and business/economic scenarios; it is flexible, evolving as the nature or number of risks, relevant business and financial strategies change; and it represents the nexus of financial and actuarial languages.

Let's consider the example of XYZ Company. The CEO and CFO of the XYZ Company want to look at all of their business risks holistically, evaluating a variety of risk/reward tradeoffs and ultimately establishing an overall risk management strategy that reflects the culture and goals of the organization.

Because XYZ is publicly held, the managers agree that the most important financial measure is earnings per share. They also establish a corporate threshold for "pain"—the risks they are unwilling to tolerate, i.e., a drop in annual earnings of more than five cents per share more than once every 20 years. They establish corresponding pain tolerances for other earnings per share threshold values. Together, these values represent the company's appetite for risk.

EXHIBIT 1

**TYPES OF RISK**


The CEO and CFO then reexamine XYZ's risks to be sure they have identified all the potential material exposures—from natural catastrophes, liability lawsuits and competitive pressures, to interest rate fluctuations, currency swings and commodity prices. Some of these risks, such as natural catastrophes and pension fund investments, operate fairly independently. However, risks such as interest rate and currency fluctuations can have a strong interrelationship. Others may depend on the individual company. For example, hurricanes may present a Florida-based construction company with the risk of property damage but also the potential for windfall profits.

Once XYZ's management understands its risk exposures and their potential for interaction, it can begin to model the extent of the company's risks under various scenarios. This model is created by reviewing past experience as well as looking forward. Examining historical experience can be valuable, but it is often only part of the story. For example, if XYZ Company produces a variety of software products and has a lively research and development department, its earnings in five years will be largely dependent on products that don't exist now. Anticipating what those products might be and what exposures they could represent is an important part of the modeling process that will involve research and development as well as the financial and operations personnel.

Catastrophe modeling is another example of looking ahead to determine potential exposures to risk. Reviewing hurricane damage over the last 10 years won't give XYZ enough information to predict its property losses for the next few years, so the company uses sophisticated natural catastrophe models that reflect its unique concentration of property values and its business interruption dynamics. Outside data sources revealing weather or

earthquake patterns might also figure into this equation.

Modeling all these exposures may seem like a formidable task. But, in actuality, the modeling is not the difficult part. The challenge is knowing enough about the nature of the exposures to make informed judgments on the specifications of the model such as likely frequency, severity and impact on operations under various business/economic scenarios. This underlying knowledge of the business is critical to effective management of XYZ in any event—whether or not it is used in a model. If this knowledge is lacking, the exercise of model-building can force the issue.

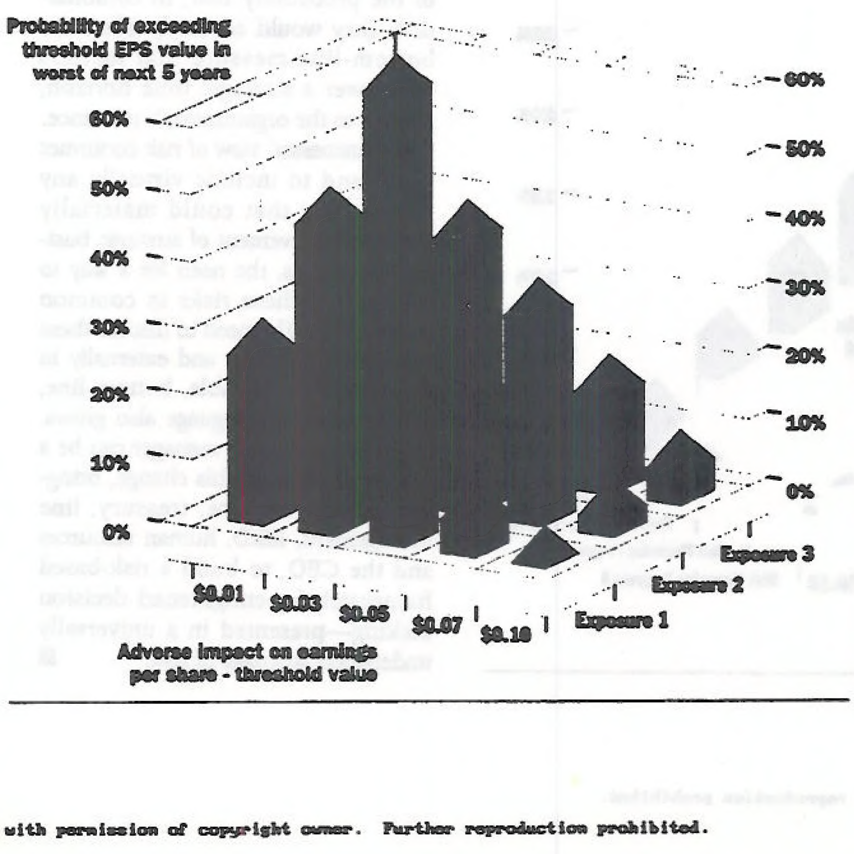
**Modeling exposures may seem like a formidable task. But, in actuality, the challenge is knowing enough about the nature of the exposures to make informed judgments on the specifications of the model.**

### Using the Model

Once a model is built, XYZ can put it to good use. Let's focus on the issue of designing a risk financing program. XYZ uses the model to simulate multiyear projections of expected costs and, more importantly, the volatility of those costs. That volatility can be expressed in terms of impact on XYZ's earnings per share. Exhibit 2 shows how that might be done. The exhibit represents the probability of each exposure—Exposure 1 might be earthquakes; Exposure 2, currency fluctuations; Exposure 3, product recall—and shows that in the worst year out of the next five, the adverse impact on earnings per share will exceed the threshold value shown. For example, the probability that an earthquake (Exposure 1) would impair XYZ's earnings by more than $.05 per share in the worst year out of the next five is 15 percent. How severe could the impact be? This can be forecasted by looking at the probabilities of exceeding higher EPS threshold values, e.g., $.07 per share (10 percent probability in this example) and $.10 per share (5 percent probability).

EXHIBIT 2

**Individual Exposure Model**




Although this information is useful, the real value of the exercise comes from allowing these various risks to interact. While a number of the exposures can be considered independent (and their simulation results simply combined), some will be interdependent. The modeling of interdependent risks is useful, however; understanding the nature of the dependent exposures is imperative to sound management.

Now XYZ can test the efficacy of various multiyear, multirisk financing programs. Exhibit 3 shows an example of this approach. Under risk financing program A (which may be a combination of high-layer catastrophe insurance and an earnings-smoothing financial product, both on a combined-risk basis and a carve-out for California earthquake), the probability that the combined impact of all exposures would result in a hit to earnings of more than $.05 per share in the worst year out of the next five is 2 percent.

After completing the exercise, XYZ discovers that by examining all its material risks on a combined basis, the risks' aggregate impact on earnings is less than if each were treated and hedged individually. In other words, XYZ's prior strategy of hedging each risk separately from the others upset the "natural order of things" and, by not allowing independent risks to offset each other over time, led to risk financing/hedging programs that were collectively costly and counterproductive.

The final piece of the picture is to create an overall representation of XYZ's risk tolerance. This is conveyed in Exhibit 4 as a ceiling that sits above the cityscape of risk. An objective of XYZ's risk financing program is to stay below this limit. The elements of this picture constitute XYZ's "riskscape." ➤

EXHIBIT 3

## Combined Exposure Model

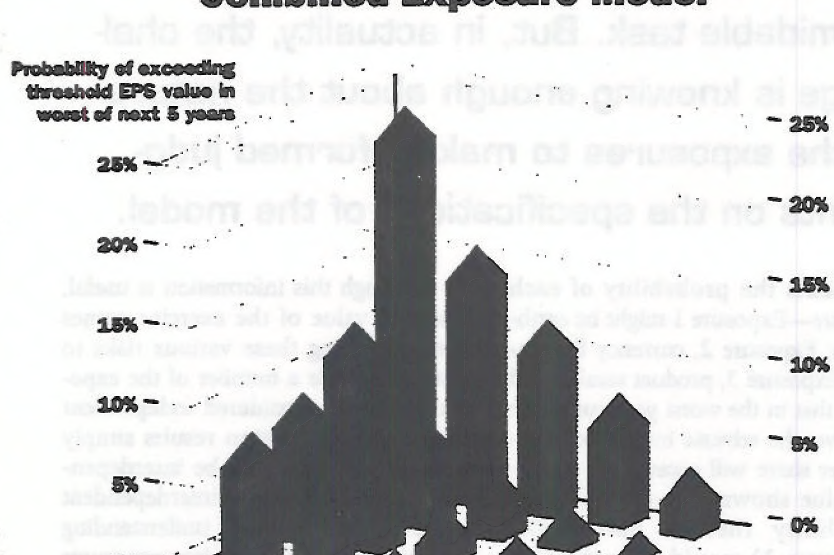
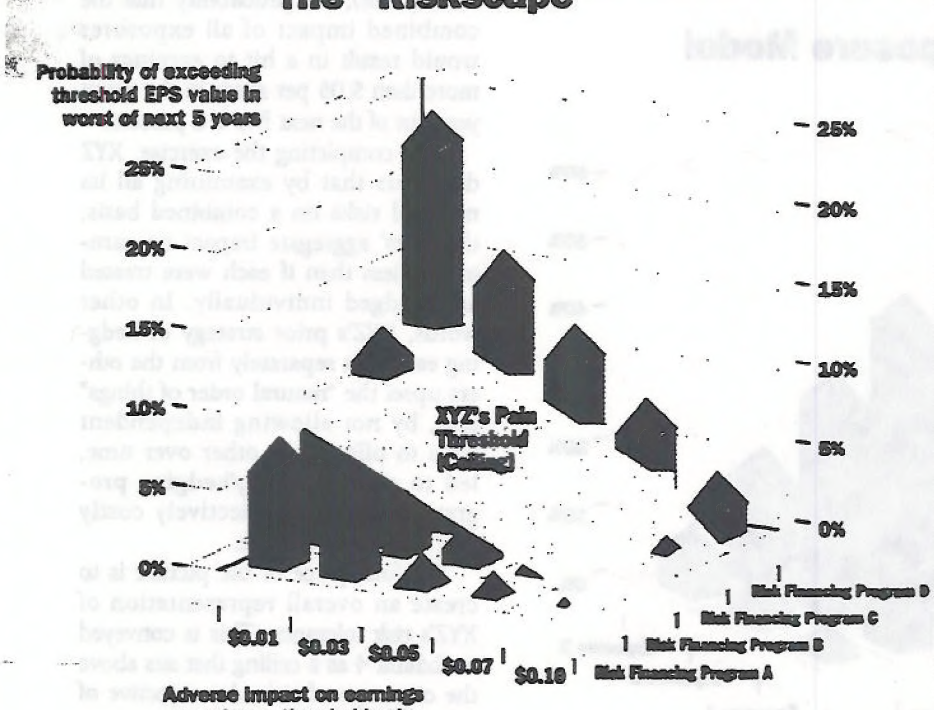



EXHIBIT 4

## The "Riskscape"




How does this modeling affect the company's focus? In the past, XYZ had made its risk financing decisions—indeed, most of its business decisions—by focusing on the reward side of the risk/reward tradeoff. Net present value analyses ruled in this context. Now management has a tool to evaluate the risk side of the equation. Risk financing alternatives can be judged, and most business decisions made, based on how they change the shape of XYZ's riskscape. This visual device has in effect become the common language for risk within XYZ. An important result has been that more effective management of the riskscape—changing the risk profile—has become the focus of risk management within XYZ.

### The Overall Picture

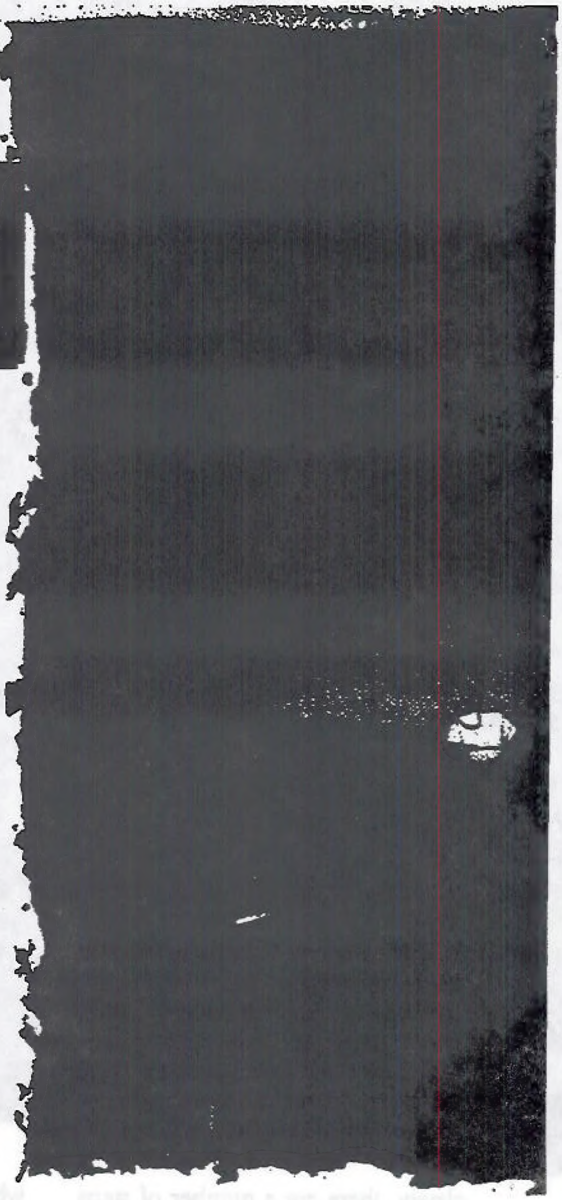
The XYZ Company is but one example. ABC Ltd. performed a similar exercise and found the most important factor to be the probability of a material interruption in annual revenue (turnover) growth in the worst year out of the next seven. DEF Co. preferred to use cashflow on a three-year rolling average basis over the next 10 years. The unifying concept of common language in all these examples is that all sources of risk to the organization are represented in terms of the probability that, in combination, they would adversely affect the bottom-line measure that matters most over a strategic time horizon, relative to the organization's tolerance.

As businesses' view of risk continues to expand to include virtually any uncertainty that could materially threaten achievement of strategic business objectives, the need for a way to understand these risks in common terms grows. The need to discuss these risks both internally and externally in clearly understandable, bottom-line, business-context language also grows. The progressive risk manager can be a natural catalyst for this change, bringing together finance, treasury, line management, R&D, human resources and the CEO, to build a risk-based framework for enlightened decision making—presented in a universally understood language of risk. ■

2



# AN Integrated Approach TO RISK MANAGEMENT



**by Todd L. Williams**

We read more and more about risk manifesting itself at otherwise-successful organizations. The stories of how the foundations shook at Orange County, Barings PLC, Proctor & Gamble, Exxon and Daiwa are the kinds of things that cause managers to lose sleep. Which company will fall victim to the next such incident? Will it be employee sabotage, product tampering or a sudden regulatory shift that causes a potentially significant reduction in earnings?

Business leaders are asking why business risks of all types, not just financial risks, are causing trouble more often today. Change, such as reengineering and downsizing, as well as the complexity of the global economy and technological advances, causes risk. And because change is occurring today at a much faster rate than ever before, it compounds the emerging risks, decreases margins of error and shrinks decision-making timeframes.

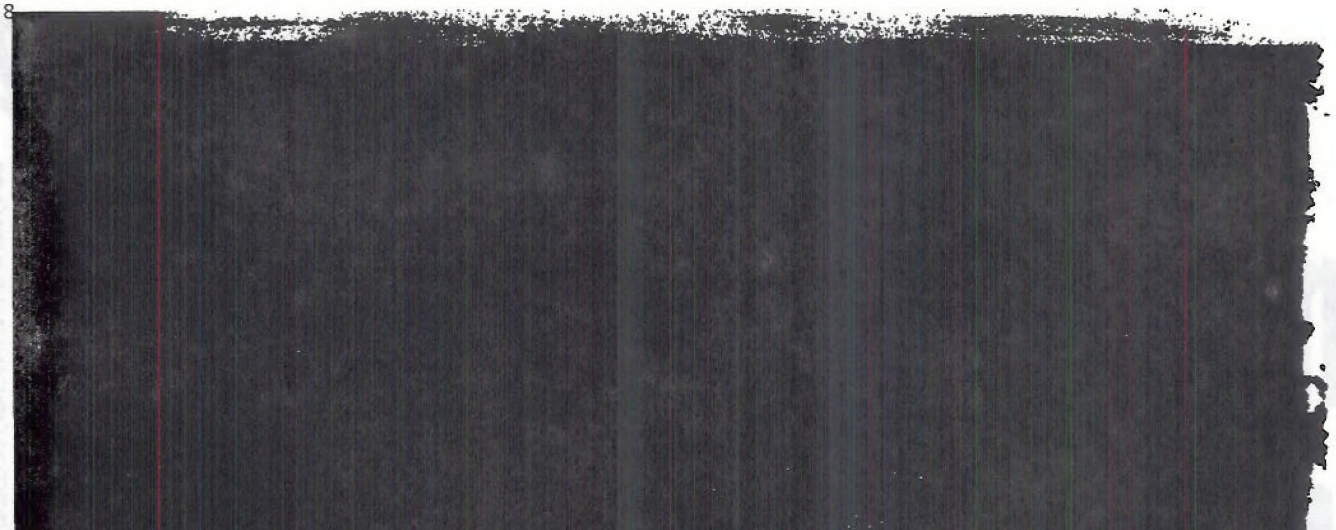
In short, the business of risk is changing and, therefore, so must the business of risk management. Although many organizations have traditionally viewed risk as residing only in neatly defined, insurable exposures, business risk lurks in all functions and processes. In this tumultuous environment, yesterday's approaches may not keep pace or accurately protect a company from risk. Business risk must be redefined as any event or action that prevents an organization from achieving its objectives—and the job of managing risk must be shared by executives in all operating disciplines.

To understand how organizations are handling business risks, Arthur Andersen and The Economist Intelligence Unit conducted a global study over the past 18 months. Written surveys, phone interviews and face-to-face discussions were utilized to gather information from senior executives about the steps organizations are taking to get a handle on their risks.

Some of the most surprising results indicated that more than half of senior executives interviewed were not highly confident that all of their relevant business risks had been identified and placed under adequate controls. In addition, 50 percent said they had reviewed their risk control policies and procedures in the last year, and almost two-thirds of these same executives stated they would be reviewing their risk assessment tools and techniques within the next three years.

While it is vital to stress that no organization can eliminate risk com-

*Todd L. Williams, CPCU, ARM, is a risk management consultant for Arthur Andersen in New York. This article is based on findings of the study* Managing Business Risks: An Integrated Approach, *prepared by The Economist Intelligence Unit and Arthur Andersen.*



pletely, there are a number of steps organizations can take to be more proactive in managing the total universe of risks they face:

**Create a process-oriented view of risk management.** Most companies are organized around traditional functions. Process structures based on activities, rather than reporting relationships, are more flexible than typical vertical structures (and thus easier to change). A process approach to risk management tears down traditional functional barriers, allowing business risk assessment and control to permeate all aspects of an organization. The risk management process can be strengthened considerably when its application involves effective cross-functional communications and coordination.

**Build an effective organizational control structure.** Responsibility for alerting senior management to relevant risks cannot sit with one person, team or department. All managers must have direct communications channels up, down and across their business units to facilitate risk identification and implementation of the appropriate responses. Traditional organizational frameworks are giving way to new models like boundaryless communications at General Electric, horizontal structures at Allied-Signal and flattened pyramids at Eastman Chemical. These new models all share a looser, information-based structure that depends more on information flows than reporting relationships.

**Develop a common business risk language.** All contributors to the business risk management equation must be able to understand one another. An organization cannot adapt to a changing environment without effective information flows across functions, divisions and reporting channels. Ian Ramsden of Analog Devices, the Massachusetts-based semiconductor company, says, "It is the commonly defined language and a routinely applied problem-

JOHN BERRY

solving process that allow my company to organize cross-functional teams rapidly in an extremely competitive marketplace."

This last step is especially relevant to those who perform the role of the traditional risk manager. Tools such as a common risk language allow us to talk directly to individuals from the boardroom to the boiler room and to explain risk control scenarios with clarity and depth.

**Expanding Risk Management**

What does research that shows the necessity for a heightened awareness of risk mean for traditional risk managers? It presents an opportunity to redefine the function's scope and activities within an organization.

The traditional risk manager's focus has been *pure risk* (defined as uncertainty as to whether a loss will occur), and it has been his or her responsibility to prevent, fund, transfer, avoid or control fortuitous losses. A focus on *speculative risk* (uncertainty as to whether a specific activity is likely to produce gain or loss) has been reserved for operational managers. Expanding the definition of risk to encompass pure and speculative risk can lead to a holistic, integrated approach to managing all business risks (as opposed to considering each type of risk independently).

This expanded definition of risk also has the potential to include a broader strategic perspective. For example, Arthur Andersen helped one client increase senior management's awareness of a probable legislative shift that would change one profit center's customer base from a small group of federal agencies to similar entities at the state level. Such a change would alter the division's marketing strategy and could ultimately change the allocation of capital to the division. This application of the risk management process to a strategic situation has traditionally been outside the risk manager's role.

Armed with a new focus and expanded scope, the risk manager's principal activities will change. The new risk manager's role will be that of a "risk process manager" or "risk champion" for the organization. The risk process manager will be challenged not only with traditional insurable risks, but also with applying the same risk management process to all business risks. Effective risk managers will start by developing a common risk language and establishing an effective control structure that promotes an integrated risk management process throughout the organization. The shifting risk management paradigm is illustrated in Exhibit 1.

Specific examples of organizations that have a fully integrated approach to managing business risks remain rare. However, this author is aware of several risk managers and organizations that appear to be on the leading edge, such as Roger Lewington of The BOC Group, Scott Lange of Microsoft Corporation, and Timothy Bunt of Prudential Insurance Company of America. These risk managers have several things in common. They use the expanded definition of risk; they each have the support and involvement of their senior management; and they are all taking a process view to manage risks. These individuals currently represent exceptions to traditional roles, but as risk increases as quickly as the rate of change, risk managers will have the chance to increase their visibility and importance within their organizations.

As risk experts, risk managers are positioned ideally to play a leading role in managing the organization's business risks as well as to increase the challenge and importance of their jobs. Adhering to a traditional function will likely cause risk managers to miss an opportunity and to see the role of managing business risks championed by someone else.

**EXHIBIT 1**

## Reinventing Risk Management

| Old Paradigm | New Paradigm |
|---|---|
| Risk management applied only to pure risk | Risk management applied to pure and speculative business risks |
| Functional approach, limited to the risk management department | Process approach transcending functions and divisions |
| Operational perspective | Operational and strategic perspective |
| Risk manager | Risk process manager or risk champion |
| Senior management support | Senior management support and involvement |
| Insurance jargon understood by a few | Common risk language understood from the boardroom to the boiler room |

**An Integrated Approach**

How does an organization apply this integrated approach to risk? One example comes from a major U.S. publishing and financial services company where senior financial management established an objective to identify, measure and assess business risks and controls. In addition, senior management wanted a high-level analysis of gaps in risk control strategies and recommendations for closing those gaps. The following is a brief description how this process was completed.

To identify strategic and operational risks, roughly 50 key execu-

tives were interviewed. At the corporate level, the CEO, president, treasurer, director of risk management and the senior executive for each function (such as human resources, legal, management information systems and others) were interviewed. Similar interviews were conducted among divisional or profit center managers.

The next step was to measure those risks and assess the existing controls. When controls for exposures that exceeded a predetermined financial threshold were examined, we found some risks with inadequate controls and some that were actually overcontrolled.

Next, Arthur Andersen presented the study's findings to the company's senior executive committee, and it responded by charging key members of senior management with responsibility for certain under-controlled risks. For this organization, the expertise required to close risk control gaps was found internally; however, in many instances, companies must search outside for the required expertise. One year later, the company is still proactively eliminating risk control gaps identified during the process.

The key success factors in this instance were expanding the definition of risk, applying a process-oriented view to risk management and enlisting the support and involvement of senior management. The company plans to perform this process every three to five years. By including pure and speculative business risk and focusing on potentially catastrophic exposures, financial executives easily gained the interest and support of senior managers at the corporate and divisional level.

EXHIBIT 2

### 10 Risk Management Warning Signs

**Not linking risk to value** - Business risk is not clearly linked to its effect on shareholder wealth.

**No effort to anticipate** - Organizational risk evaluations do not keep pace with the ever-changing global business environment.

**Ineffective strategic control** - Environmental and other assumptions are rarely evaluated until they become obvious that a strategy is not working.

**No business risk policy** - There is no written requirement for managers to periodically assess business risk and performance of risk controls.

**Not a priority** - Management does not make clear that ongoing risk assessment and control is a high priority. Consequently, employees and managers do not consider it important.

**No integrated risk assessment framework** - There is no systematic approach for identifying, measuring and monitoring business risks.

**Fragmentation** - Functions and departments operate independently as separate silos. Operating managers think business risk management is someone else's job.

**Narrow focus** - Finance focuses only on internal accounting risk control. Treasury focuses only on controlling a few financial risks.

**Poor risk communications** - Periodic, sporadic internal communications about risk assessment and control focus only on damage control.

**Too little, too late** - Specialized expertise for assessing and controlling complex risks is used reactively to put out fires. After-the-fact evaluations by internal auditors are an integral part of the front line of defense.

Source: "Managing Business Risks: An Integrated Approach," The Economist Intelligence Unit and Arthur Andersen

#### Understanding Risk

We have performed the same type of process with numerous other clients and uncovered an array of business risks. One common exposure stems from the redistribution of management information, such as marketing databases, accounts receivable, accounts payable and other critical functions, from centralized data centers to networks at local operations. In many cases, a company's primary data center has strong risk controls, such as a disaster plan, security measures and off-site data storage. However, as technology advances, we have found operations that have moved management information away from a centralized data center without installing the appropriate risk control measures—leaving the division exposed to risk.

Another type of business risk includes legal precedents that could dramatically affect the way a company does business and its financial results. A current example is the recent settlement by one company in the tobacco industry, which could influence similar litigation faced by other companies. Other scenarios include legal decisions that can influence how a particular product can be marketed or that change the prevailing standard of care or expense needed to manufacture a product.

Other common risks include damage to a company's reputation, regulatory or legislative changes, elimination of barriers to entry, financial

**As the names of companies that have underestimated their exposures are plastered across news reports, it is no longer enough to ask the old questions such as "Are my risks covered?".**

risks and catastrophic loss. Warning signs that risk management efforts are not controlling these exposures are listed in Exhibit 2.

### Changing Perceptions

Risk control measures are typically perceived as constraints and boundaries restricting the activities of operational managers. There is more to this analysis, though. Attempting to assess and control all relevant business risks proactively not only helps to prevent adverse events from happening, it also allows favorable results to develop. Although evaluating where an organization's risks lie will probably indicate a need to allocate resources to specific threats, this search is likely to uncover some risks that are overcontrolled. When this occurs, managers can reallocate scarce resources to match the organization's objectives more effectively.

Risk and risk management may be the words of the day, but business risks and the problems they can cause are not going away. Risk has always been (and will always be) here. Today's difference is that we are starting to notice the effects of risk more often. The good news is the business world is responding. Chief risk officers are being named by major organizations to try to allay these fears.

As the names of companies that have underestimated their exposures are plastered across news reports, it is no longer enough to ask the old questions such as "Are my risks covered?". Risk managers today must go at least one step further by asking "How do I know? Am I sure?".

By broadening the current definitions of risk and risk management and taking more proactive steps to identify risks and their sources, you can better fulfill your corporate responsibility as a senior executive, operations manager or risk manager. You also change from your current role, whatever it may be, to risk manager. Shouldn't we all be risk managers in this risky world?

CHAPTER 3

# RISK AND UTILITY: ECONOMIC CONCEPTS AND SIMPLE DECISION RULES

Risk is present when the outcome of some defined activity is not known. Given the financial or economic approach of this text, the outcomes with which we are concerned can be directly or indirectly measured in money terms. *Risk* refers to the quality of variation in the range of possible outcomes; the greater the potential variation, the greater the risk. In the economic sense, risk does not refer to the adverse quality of some outcomes (losses instead of profits), but rather to the lack of knowledge about which of several outcomes may prevail. Risk is implied by our inability to predict the future. If risk was not present, many of our decisions would be trivial: We would simply choose that course of action which has the highest certain payoff. Under such circumstances, risk management would be a simple clerical task and this book would serve no purpose. However, when chance intervenes in the selection of outcomes, decision making becomes simultaneously more complex and more personal. Even though we may reasonably assume that everyone prefers more money to less, people differ in how they respond to risk. The differences in personal preferences to risk may relate to our personalities and to our economic circumstances. Accordingly, if economics is to provide us with a useful framework for making decisions under conditions of risk, it must help us to process and compare the potential outcomes on terms dictated by our personal preferences and circumstances.

Our task in this chapter is to specify simple decision rules that will help us to come to terms with the nature of risk and its effect on decision making. The so-called expected-utility hypothesis will be used to analyze simple risk management decisions—notably insurance decisions. In addition, the basic ideas will be useful in making preliminary statements concerning loss prevention and in analyzing certain features of the insurance market. The concepts to be discussed in this chapter are preliminary.



3

At first approximation, we may think of the simple risk-management decisions discussed here as being faced by an individual rather than by a corporation. The individual must make decisions on financing or take steps to prevent events that might prove catastrophic to him or her. When we come to look at corporations later, the separation of ownership and control will call for considerable modification of the decision procedure. Ownership of a corporation is represented by financial claims (shares of stock). We must then determine how the risky prospects facing a firm affect the ownership claims and how risk-management decisions can best protect the welfare of a firm's owners. This will require some development of statistical measures of risk, an understanding of the process of diversification, and some rudimentary awareness of how capital markets function. These are the subjects of the following chapters. For now, we will concentrate on an individual who must make decisions concerning risky events that may profoundly affect his or her personal welfare.

## DERIVATION OF DECISION RULES

### Simple Risky Prospects

A convenient starting place for our exploration of risky decision rules is an analysis of choices in which no risk is present. The decisions to be made in this uncomplex and safe world turn out to be trivially simple if we assume that outcomes can be specified in money terms. Of course, some events bring payoffs that are not directly specified in money values, but we assume that money equivalents can be assigned by the decision maker. For example, consider a choice between action A and action B. Action A is simply "do nothing" and has a zero payoff. Action B involves working for 10 hours for a fee of $200. To make a comparison, the decision maker must place a money-equivalent value for the labor involved in action B. If he decides that he is indifferent between each hour's work and receiving $18, then 10 hours' effort has a negative money value of $180. The comparison between actions A and B is now presented in the following terms:

| Action | | Payoff ($) |
|---|---|---|
| A (Do nothing) | | $ 0 |
| B (Work 10 hours) | Fee | $200 |
| | Value of effort | −180 |
| | Net value | $ 20 |

As long as the decision maker is confident that he has correctly valued his time and effort at $18 per hour, the decision is quite mechanical. Action B has a higher payoff than action A and should be preferred.

To stretch our assumption of a certain world, consider a "gamble" between two people who have the unnerving facility to predict the future. The "gamble" involves the flip of a coin, the outcome of which, for us mortals, is certainly a risky activity, but which, for our prophets, is never in doubt. Prophet A will pay prophet B $10 if the coin turns up heads; otherwise prophet B will pay prophet A $10. Since both know that the coin will turn up tails, a meaningful transaction will never take place; prophet B will never agree. Possibly prophet A might be willing to pay prophet B $10 to induce him to make the "bet," but the activity is pointless because both break even. If we descend from Mount Olympus, the transaction will have quite different implications, since the outcome from tossing a coin is quite risky. Neither party will, in reality, know the outcome in advance, and the gamble therefore exposes both parties to a risky future. The payoff to the parties depends on chance, as follows:



| Decision | Payoff |
|---|---|
| A's Decision: | |
| 1. Gamble | $10 or −10 |
| 2. Do not gamble | $0 |
| B's Decision: | |
| 1. Gamble | $10 or −10 |
| 2. Do not gamble | $0 |

To help A and B decide whether to gamble with each other, they might each make use of further information they have at their disposal. Neither knows whether the coin will land heads or tails, but each, believing the coin to be unbiased, assumes that the coin is equally likely to land on heads and tails. In other words, A and B believe that the probabilities of heads and tails are each one-half. We will represent this information, together with the payoffs, in the following format (the alternatives available to each decision maker are prospects $V$, that is, gamble, and $W$, that is, do not gamble):

$$V \text{ (gamble)} = \begin{cases} \$10 & 0.5 \text{ probability} \\ -\$10 & 0.5 \text{ probability} \end{cases}$$

$$W \text{ (do not gamble)} = \begin{cases} \$0 & 0.5 \text{ probability} \\ \$0 & 0.5 \text{ probability} \end{cases}$$

The dollar value on each line represents the payoff, and the associated probability follows.

In this choice, risk is present in the spread of values for the first alternative, but it is absent from the second alternative. With the gamble, the outcome cannot be predicted; it can take a range of values. Risk is inherent in this range or spread. If no gamble is selected, the outcome is the same, regardless of some chance event such as the toss of a coin. However, risk is a relative term. Consider the preceding gamble $V$ along with the following:

$$X = \begin{cases} \$10 & 0.99 \text{ probability} \\ -\$10 & 0.01 \text{ probability} \end{cases} \qquad Y = \begin{cases} \$10 & 0.01 \text{ probability} \\ -\$10 & 0.99 \text{ probability} \end{cases}$$

$$Z = \begin{cases} \$20 & 0.5 \text{ probability} \\ -\$20 & 0.5 \text{ probability} \end{cases}$$

$X$ has the same range of outcomes as $V$, but the probabilities have changed. Certainly, the odds have been altered in favor of the gambler, but the degree of predictability has also changed. Winning comes close to being a "sure thing." Thus chance has less opportunity to be capricious. Although the range of possible outcomes remains unchanged, the degree of "variability" has changed significantly. $Y$ simply reverses the probabilities from $X$. Now the odds are biased against winning, and undoubtedly, this reduces the attractiveness of gambling. However, the degree of "predictability," or variation in outcomes, remains similar to that for $X$. Our concept of risk will take account of probabilities. Comparing $Z$ with $V$, the probabilities are the same, but the range of outcomes has increased. The increase in the range of outcomes contributes to an increase in risk. Thus *risk*, in the economic sense, is a quality that reflects both the range of possible outcomes and the distribution of respective probabilities for each of the outcomes.

### The Expected-Value Rule

Let us now set up a decision rule that can help us to choose between risky alternatives. Consider, for example, a choice between the following (outcomes in dollar values, followed by probability):

$$A = \begin{cases} 10 & 0.5 \\ 10 & 0.5 \end{cases} \qquad D = \begin{cases} 0 & 0.4 \\ 20 & 0.6 \end{cases}$$
$$B = \begin{cases} 0 & 0.5 \\ 20 & 0.5 \end{cases} \qquad E = \begin{cases} 0 & 0.6 \\ 20 & 0.4 \end{cases}$$
$$C = \begin{cases} 5 & 0.5 \\ 15 & 0.5 \end{cases} \qquad F = \begin{cases} 1 & 0.5 \\ 21 & 0.5 \end{cases}$$

Each alternative will be called a *prospect;* outcomes $B$, $C$, $D$, $E$, and $F$ may be further described as "risky" prospects. Outcome $A$ is clearly a nonrisky prospect.

The decision rule will focus on the outcome that may be expected to prevail "on average." More precisely, we will select the alternative with the highest expected value $EV$, defined as follows:

$$EV_j = \sum_i P_i X_i \qquad (3\text{-}1)$$

where $EV_j$ = expected value of the prospect $j$
$P_i$ = probability of outcome $X_i$
$X_i$ = outcome in money value

Thus, for alternative $B$, the expected value is

$$EV_B = 0.5(\$0) + 0.5(\$20) = \$10$$



One interpretation of this value is as follows. If an individual accepted a large number of such risky prospects, he or she would win \$20 on some and nothing on others. The average outcome would be very close to \$10 per prospect if the number of prospects was sufficiently large.[1] Thus \$10 is the long-run average value for the risky prospect. If the individual paid \$10 for each bet of type $B$, he or she would break even in the long run.

Using this expected-value rule, a single value can be assigned to each prospect, and this, hopefully, will aid in decision making. The respective values are

$$EV_A = EV_B = EV_C = \$10$$
$$EV_D = \$12$$
$$EV_E = \$\ 8$$
$$EV_F = \$11.5$$

Consequently, our ranking should be

$$D > F > (A \text{ or } B \text{ or } C) > E$$

If this rule works successfully, testing a group of decision makers should confirm the ranking. Testing on a class of graduate and undergraduate students reveals some problems. Preference of $D$ over $E$ appears to be virtually unanimous for all subjects. So far so good; this is exactly what the rule predicts. However, preferences tend to appear between $A$, $B$, and $C$, and according to the rule, they should rank equally. Most students prefer $A$ to $B$ (or $C$) and $C$ to $B$. One stubborn approach to such problems is to assert that students do not know what is good for them and that we should continue to believe the rule. However, the rule must *serve* preferences, not be a master to them. Other problem rankings tend to arise. The expected-value rule asserts that $F$ should be preferred to $A$, but many students (not all) prefer $A$. Several other choices do not correspond with the expected-value rule. Such experiments are not laboratory-controlled, and differences between group responses appear. However, such simple tests do reveal that individuals have strong preferences that cannot obediently be classified by such a simple rule. Having preferences different from those predicted by the expected-value rule does not imply irrationality.

### The St. Petersburg Paradox

A more dramatic illustration of how such a rule may break down was formulated some two centuries ago by Daniel Bernoulli. Bernoulli was an eighteenth-century Swiss mathematician and physicist and one-time professor of mathematics at the Russian

[1] This also assumes that successive outcomes are independent. This will be discussed in detail later.

Academy in Saint Petersburg. Bernoulli posed the following problem, which has come to be known as the *St. Petersburg paradox:*

> Consider a game in which one player flips a fair coin. If the coin reveals heads, the player will pay the other player 2 dollars (or ducats or whatever), and the game is over. If the coin lands on tails, it is tossed again. If the second toss lands heads, the first player pays $(\$2)^2$ dollars to the second, and the game is over. If the second toss lands tails, the coin is flipped again. Thus the game continues until the first heads appears, and then the first player pays the second $(\$2)^n$, where $n$ signifies the number of tosses required to reveal the first heads. Since a head will turn up eventually, the second player wants a long preceding run of tails. The problem is: How much will the second player be willing to pay to enter the game? Of course, individuals vary in their responses, but seldom will anyone be willing to pay more than \$10 to enter such a game.
>
> The nature of the paradox can be seen by considering the expected value of such a game. We will represent the risky prospect in the following form, showing the sequence of tails and heads required to produce each outcome:

| | Outcome | Probability | Sequence required to produce outcome |
|---|---|---|---|
| | \$2 | $^1/_2$ | H |
| | $2^2$ | $^1/_2{}^2$ | TH |
| | $2^3$ | $^1/_2{}^3$ | TTH |
| St. Petersburg Paradox = | . | . | . |
| | . | . | . |
| | . | . | . |
| | $2^n$ | $^1/_2{}^n$ | TT · · · [($n$ − 1 times)] · · · H |

> The expected value of the game is
>
> $$\sum P_iX_i = 2(^1/_2) + 2^2(^1/_2)^2 + 2^3(^1/_2)^3 + \cdots + 2^n(^1/_2)^n + \cdots$$
> $$= 1 + 1 + 1 + \cdots$$
> $$= \infty$$
>
> Why is it that people will typically only pay a few dollars to participate in a game that has an infinite value? Clearly, the expected-value rule has broken down dramatically.

What causes the expected-value rule to fail in the simple-choice experiments given to the class and in the St. Petersburg paradox? The answer is that risk is ignored. Returning to the simple prospects, \$10 with certainty (prospect $A$) is clearly very different from a fifty-fifty gamble for \$20 or nothing (prospect $B$). The general preference for $A$ over $B$ and the small entry price to the St. Petersburg game indicate that the quality of variability in outcomes is an adverse factor for the decision maker given equality in expected values. Most people prefer choices with less risk rather than more. This does not mean that people will never choose risk when it can be avoided. The expected value of one prospect may be so advantageous in comparison with a competing



prospect that aversion to risk is overcome. It is easy to construct such choices. Consider the following:

$$G = \begin{cases} \$10 & 0.5 \text{ probability} \\ \$10 & 0.5 \text{ probability} \end{cases} \qquad H = \begin{cases} \$9 & 0.5 \text{ probability} \\ \$20 & 0.5 \text{ probability} \end{cases}$$

$G$ may be seen as the entry price to gamble $H$; by not gambling, the individual retains \$10 with certainty. If I were to make an offer of accepting such gambles, I imagine that I would be killed in the rush. I hasten to add that idle speculation does not constitute such an offer. The next question therefore is: How can attitudes toward risk be considered in an appropriate decision rule? One useful answer lies with the concept of utility.

### The Expected-Utility Rule

There are two related ways of trying to resolve the St. Petersburg paradox. One approach, suggested by Bernoulli, is to place some weighting on money outcomes; later, this approach led to the concept of utility of income or wealth. A second approach is to recognize explicitly that risk per se has a cost and to embody this cost as a parameter in the decision. This second approach will be examined later when we discuss the mean-variance rule. We will now present the utility approach and show that this embodies a welfare measurement of risk.

*Utility* is a scale of measurement of the satisfaction derived from some economic good, particularly income or wealth. To represent utility, let us make some fairly innocuous assumptions. First, we will order levels of satisfaction with the statement that "more wealth is preferred to less wealth." Second, we will make a further statement that has the effect of ordering differences in wealth. The incremental utility, or satisfaction, from unit increases in wealth decreases as wealth increases. Thus, if I am poor, an addition of \$1000 to my wealth will make a considerable impact on my level of welfare. If I am rich, the extra \$1000 will still increase my satisfaction, but only marginally. These concepts are known in economics as the *law of diminishing marginal utility*, and the law is represented by utility function $OA$ shown in Figure 3-1. Of course, different individuals will scale different levels of wealth in different ways. Thus two other utility functions, $OB$ and $OC$, are also shown in the figure. The positions of these functions have little significance, because they apply to different people. We are unable to make direct comparisons of different individuals' capacities for enjoying wealth. However, some forms of comparison can be made. The function $OB$ seems to be less concave than $OA$, whereas $OC$ appears to change its slope more rapidly than $OA$ as wealth increases. In other words, individual B (with curve $OB$) appears to value increases in wealth almost as highly at high levels of wealth as at low levels of wealth. In contrast, the satisfaction curve $OC$ of individual C flattens out very quickly, indicating that individual C's capacity for enjoying additional wealth tails off quickly as wealth increases. These features, as we shall see, reveal differences in attitudes toward risk.

What is the connection between utility and risk? Consider a simple gamble. The

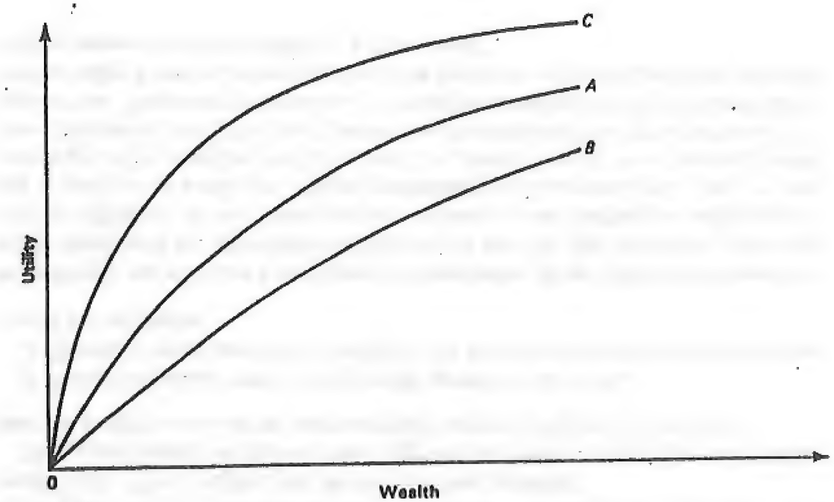



FIGURE 3-1

choices are (1) to keep our current wealth and not gamble or (2) to enter an "equal odds" gamble in which we win or lose \$10. Put another way, the gamble substitutes for our current wealth of \$10, a fifty-fifty chance of wealth \$0 or wealth \$20. Should we enter the gamble? Let us represent the choices on the wealth scale of Figure 3-2, which also shows a utility function conforming to the preceding assumptions. If the gamble is not undertaken, wealth is \$10, which has a utility value shown on the vertical axis as $U$(\$10). For an income of \$10, the slope of the utility curve is neither very steep nor very flat, showing an intermediate value on changes in wealth. Now consider a very small gamble in which we either win or lose \$1. If we win, wealth increases to \$11, which, naturally, has a higher utility than \$10. If we lose, our wealth falls to \$9, which has a lower utility than \$10. However, notice the respective sizes of these prospective changes in utility. Because the utility curve becomes flatter as wealth increases, the loss of utility from the fall in wealth exceeds the increase in utility from an equal gain in wealth. This differential valuation of the gain or loss from the gamble causes a bias against taking such a risky decision. Now consider the larger gamble. Losing reduces wealth drastically to position \$0, where marginal dollars have a very high value, as indicated by the steep slope of the utility curve at this level of wealth. However, winning takes us to \$20, where the utility curve is pretty flat, showing that marginal dollars have relatively low utility value. If the utility curve is of the form shown, gambles represent a trade in which the individual sacrifices dollars that are highly valued (because they coincide with low wealth) in exchange for contingent dollars that have a low utility value (because winning will transport the individual to a higher wealth level, where the marginal utility of wealth is low). At face value, this prospect does not appear very attractive. Diminishing marginal utility creates a bias against gambling.



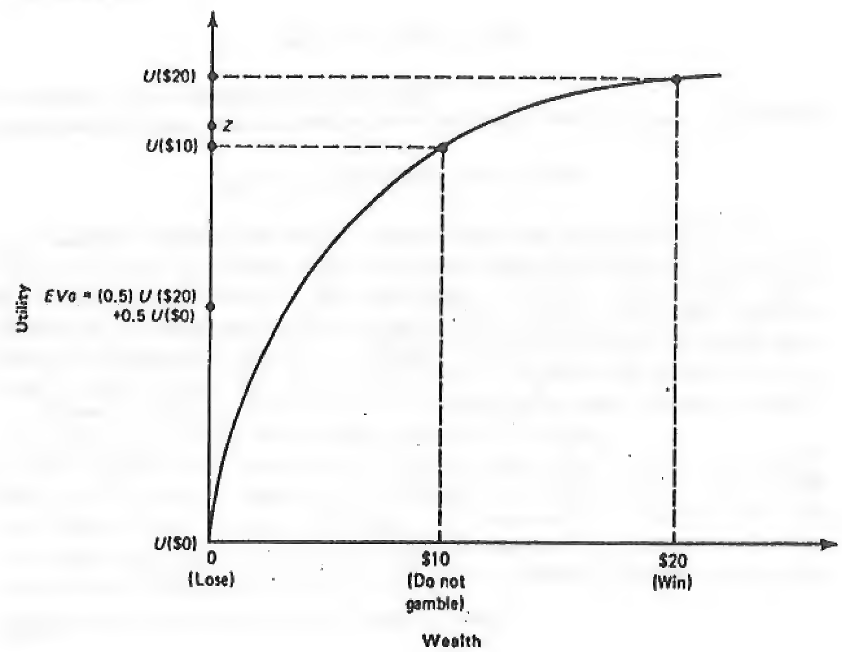
The relationship between utility and risk is now presented in a more precise fashion. The *expected-utility rule* represents a revised decision rule that substitutes utility values for money values in risky prospects in order to select among risky alternatives. The rule was developed by John von Neumann and Oskar Morgenstern (1944) and has proved to be a very powerful technique for analyzing risky choices. The rule is based on a set of formal axioms that will not be examined here; however, we will show that it can be used to provide some very useful preliminary insights into risk-management problems.

An expected utility $EU$ is assigned to a risky prospect in accordance with the following expression:

$$EU_j = \sum_i P_i U(X_i) \tag{3-2}$$

where $EU_j$ = expected utility of prospect $j$
$P_i$ = probability of outcome $X_i$
$U(X_i)$ = utility value of outcome $X_i$ derived from the individual's utility function

FIGURE 3-2



Notice how similar this rule is to the expected-value rule. Whereas *expected value* is the weighted average of the money outcomes, *expected utility* is the weighted average of outcomes when those outcomes are expressed in utility values.

The application of the rule is illustrated in Figure 3-2. The utility values of winning and losing are shown on the vertical axis as $U(\$20)$ and $U(\$0)$, respectively. Since the odds on winning and losing are each 0.5, the weighted average of the two utility values is the halfway point; that is, the expected utility of the gamble $EU_G$ is

$$EU_G = 0.5U(\$20) + 0.5U(\$0)$$

The position is shown on the vertical axis. Now consider the alternative choice, which is to avoid the gamble and keep current wealth with certainty (i.e., a probability of unity). Writing the expected utility from not gambling as $EU_{NG}$, we have

$$EU_{NG} = 1.0U(\$10) = U(\$10)$$

Clearly, the expected utility from not gambling is higher than that from gambling. Therefore, according to the decision rule, the gamble should be avoided.

There is no indication from this analysis that one should never gamble, but it does imply that gambling should not be undertaken at such odds. The preceding example reveals fair odds; that is, the entry price is equal to the expected payoff. There is an entry price for the gamble of zero and the expected payoff is also zero, as shown:

$$EV = 0.5(-\$10) + 0.5(\$10) = 0$$

It is possible to play with the odds until the gamble becomes attractive. For example, at 0.9 probability of winning, the expected utility is recalculated at point $Z$ on the vertical axis. This is higher than the utility of not gambling.

From these results, we can make some general statements that apply to an individual who has a utility curve of the concave shape shown in Figures 3-1 and 3-2:

1 These individuals would not rationally gamble at fair odds.

2 The odds would have to be loaded in the individual's favor in order to induce him or her to gamble.

Accordingly, we will state a very important conclusion. If an individual has a concave utility function of the form shown in Figures 3-1 and 3-2, that individual is averse to risk by definition. As we shall soon see, aversion to risk implies the willingness to pay a premium to avoid risk. Before proceeding, the interested reader may try constructing a utility function that is convex, i.e., slopes upward at an increasing rather than a decreasing rate. The same gamble can be represented and the expected utilities recalculated. From such an exercise, it should be concluded that an individual with a convex utility function derives positive value from risk and would gamble at fair odds. Such a person may be described as a "risk lover."



## INSURANCE AND THE EXPECTED-UTILITY RULE

An insurance policy has the opposite risk effect to a gamble. A gamble involves the sacrifice of certain wealth in order to acquire the possibility of an increase in wealth. An insurance policy involves the sacrifice of certain wealth in order to avoid the possibility of a loss of wealth. With a gamble, one pays to acquire risk; with an insurance policy, one pays to avoid risk. The insurance strategy can be valued using the expected-utility rule in an identical manner as the gamble.

Suppose you wish to insure your home. To simplify the issue, your total wealth is \$120, of which \$100 is the value of your home. The house may or may not burn down, but assuming any loss to be a total loss, a fire would reduce your wealth from \$120 to \$20. The probability of a fire is 0.25 (this does not reflect pathologic tendencies in your children; it is simply a convenient value for purposes of illustration). Since your final wealth will be either \$120 or \$20, with respective probabilities of 0.75 and 0.25, you can represent your utility in contemplating this insecure prospect as

$$EU_{NI} = 0.75U(\$120) + 0.25U(\$20)$$

where the subscript *NI* indicates that no insurance has been purchased. What if insurance is available? The expected value of the loss is

$$EV = 0.25(100) = \$25$$

Ignoring transaction costs, an insurer charging such a premium would break even if it held a large portfolio of such policies. This premium could be called a *fair premium* or an *actuarially fair premium*, denoting that the premium is equal to the expected value of loss (sometimes called the *actuarial value* of the policy). The term *fair* is not construed in a normative sense, rather it is simply a reference point.

If you are offered insurance at a fair premium, in this case \$25, should you insure? Consider Figure 3-3, which shows a utility curve revealing aversion to risk. The horizontal axis shows possible values of terminal wealth. The positions \$20 and \$120 represent possible wealth positions if you do not insure, and the respective utility values are revealed on the vertical axis. Insurance at a premium of \$25 will remove the financial impact of a chance fire, but you must sacrifice \$25 of certain wealth to pay the premium. Thus, with insurance, my wealth position will be \$120 − \$25 = \$95 with a corresponding probability of unity. The expected-utility values for insurance $EU_I$ and no insurance $EU_{NI}$ are constructed as follows:

$$EU_I = 1.0U(\$95) = U(\$95)$$
$$EU_{NI} = 0.25U(\$20) + 0.75U(\$120)$$

Both positions are clearly shown on the vertical axis of Figure 3-3. As shown, $EU_I$ is greater than $EU_{NI}$, suggesting that insurance should be purchased. Is this ranking accidental? In fact, it is not.

A geometric technique to find the expected utility of two wealth levels is shown in Figure 3-3. A chord $AB$ is drawn from the two positions on the utility curve corresponding to the alternative wealth levels. The position at which this chord intercepts the expected money value of wealth (position $C$) traces out expected utility. The reason this trick works is that the same probability weights are used for expected value and expected utility. Clearly, if the utility curve is concave from below, point $C$, which traces out the expected utility of not insuring, will always be below point $D$, which identifies the expected utility of insuring. Thus we can state the following preliminary ideas:

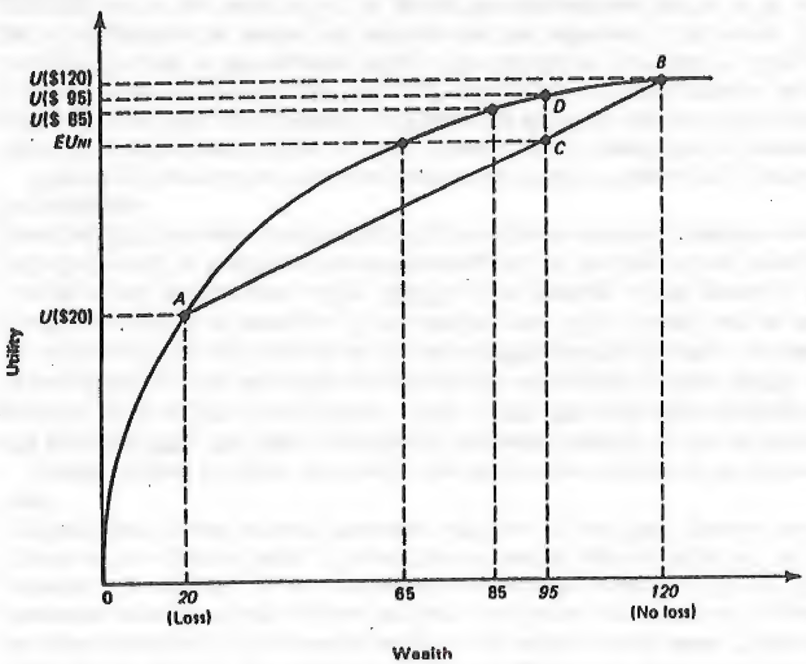
1 A concave utility curve implies aversion to risk, in confirmation of our earlier result.

2 A rational, risk-averse individual will choose to insure if the premium is equal to the expected value of loss.

This second statement is very important. It is often known as the *Bernoulli principle*, and it provides a powerful rationale for insurance and will help condition our thinking for more complex corporate risk-management problems.

Before introducing more realistic problems, let us dwell for a little while longer on risk aversion and stress the equivalence between the shape of the utility function

FIGURE 3-3





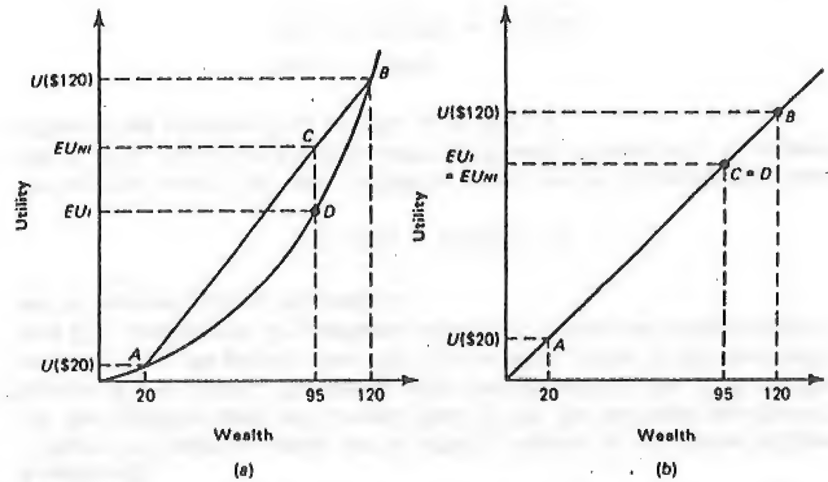




FIGURE 3-4

and attitude toward risk by considering exceptional cases. Figure 3-4($a$) shows a utility function that is convex, implying that, for this individual, the marginal valuation of wealth increases as wealth increases (the more money she gets, the more money she wants). The attitude toward insurance is quite different from that revealed by a concave utility function. Since the marginal valuation of wealth is high when we pay the premium but low when we receive a loss settlement, insurance is unattractive. The same geometric construction is undertaken, using the same notation as in Figure 3-3, to show the respective positions of $EU_I$ and $EU_{NI}$. Notice that $EU_{NI}$, derived from position $C$, is now higher than $EU_I$, from position $D$, showing that insurance should not be purchased. It takes little imagination to show that if the utility curve is linear, as shown in Figure 3-4($b$), the individual is neither averse to risk nor does she like risk; she is risk-neutral. Now the expected utility from insuring is exactly equal to that from not insuring. The risk-neutral individual will choose purely on the basis of expected values of alternative risky prospects; the risk per se has no value.

## SOME RISK-MANAGEMENT PROPOSITIONS FOR INDIVIDUALS

### Insurance with Premium Loadings

The Bernoulli principle provides a useful starting point for a treatment of risky decision making. It reveals that risk has a cost and that this cost is neglected in a decision rule that uses expected value alone. However, for practical risk management, the Bernoulli principle addresses a problem that is not very interesting. Seldom, if ever, is an insurance premium equal to the expected or actuarial value of the policy payment. Even if the insurance market is highly competitive, such that all surplus profit is

removed from insurers, the transaction costs of writing insurance must be covered and the capital employed in the insurance industry must make a normal return. Typically, transaction costs of writing insurance are high, both for the insurer and for the agent or broker. Consequently, the insurance premium usually ranges between 125 and 175% or more of the expected value of losses. Can we use the expected-utility rule to say anything about optimal insurance decisions with these, more realistic, premium structures?

Consider Figure 3-3 again, but suppose that the premium charged by the insurer is \$35 instead of \$25. The effect of purchasing insurance would be to give the insured a certain wealth of \$85 (that is, \$120 − \$35 = \$85). The utility from insuring can be ascertained by using the utility curve to plot the corresponding value, $U(\$85)$, on the vertical axis. Clearly, this value lies between $U(\$95)$ and $EU_{NI}$; therefore, insurance is still preferred to no insurance. If the premium were \$55, it would turn out that insuring would give the same utility, $U(\$65)$, as not insuring. At this premium, the individual would be indifferent between insuring and not insuring. At any premium lower than \$55, insurance would be desirable; at any greater premium, insurance would be undesirable.

This example reveals that a risk-averse individual would be willing to pay a premium above the expected value of loss in order to remove risk by purchasing an insurance policy. The maximum an individual is willing to pay above the expected value of loss is known as the *risk premium* (not to be confused with the *insurance premium*, which is simply the price of the insurance policy). The fact that an individual is willing to pay a risk premium to remove risk indicates that the individual is risk-averse. The more risk averse the individual is, the greater the risk premium that he or she will pay. Thus the risk premium provides a measure of an individual's attitude toward risk.

Risk premiums will differ among individuals, and it is not possible to offer a general theorem that states precisely the conditions under which insurance will be purchased. The most general statement that can be made at this point is that *a risk-averse individual is willing to incur a risk premium to remove the riskiness of his or her wealth prospect by means of an insurance policy.* In reality, the insurance premiums charged by an insurance company may or may not be too high to induce any given individual to purchase insurance. Conceptually, we can state that if the insurance premium satisfies the following condition, individual i will purchase insurance:

$$P - E(L) < RP_i$$

where $P$ = the insurance premium
$E(L)$ = the expected value of loss
$RP_i$ = the risk premium for individual $i$

The left-hand side of the inequality represents the portion of the insurance premium allocated to transaction costs and insurer's profit. This is known as the *premium loading* or *markup*. If we know the loss distribution and the insurance premium, we can calculate the left-hand side of this condition, but the right-hand side is a subjective



value that will differ among individuals. At any given premium loading, some people will purchase insurance and others will not. The best we can say is that, other things being equal, the higher the premium loading, the smaller the number of people who will wish to buy insurance.

From this discussion it seems that the expected-utility rule can help to discipline and organize our thinking about the purchase of insurance or similar risky decisions. We can use the rule to make general statements about the decision process and the tradeoffs an individual has to make. However, to correctly predict the actual decision made by an individual requires knowledge of that individual's utility function—not just its general shape (such as concave), but its exact form. We can never look inside people's heads, nor are people usually able to articulate and quantify a concept as abstract as utility. Despite these operational limitations, the rule does yield important insights. It focuses attention on the types of costs and benefits that are involved in the decision, and it will often enable us to narrow down the set of alternative choices so that we focus only on those solutions which are compatible with our general attitude toward risk. Some further general propositions are now appropriate.

### Gambling at Unfair Odds

The preceding proposition asserted that a risk averter would be willing to pay a risk premium (i.e., pay an insurance premium in excess of the expected value of loss) in order to purchase an insurance policy. The risk premium is really the price to the insured for converting a risky wealth prospect into a riskless wealth prospect. With this reasoning, it would appear that a risk averter would want to be paid a risk premium in order to induce him or her to voluntarily accept risk. Put another way, a risk-averse individual would have to be offered odds that were in his or her favor in order to induce him or her to gamble. First, we will demonstrate this proposition; then we will try to resolve the difficulty it creates for us in explaining why people actually gamble at *unfair* odds.

Figure 3-5 depicts a gamble that is, initially, identical to that shown in Figure 3-2. The individual starts with a wealth level of \$10. The individual can choose to keep the \$10 or gamble. The gamble might involve rolling a die. If the resulting number is even, the gambler wins \$10; if the resulting number is odd, the gambler loses \$10. Assuming the die is unbiased, the odds of winning and losing are each 0.5 and the expected value of the gamble is

$$0.5(-\$10) + 0.5(\$10) = 0$$

The expected wealth is the same for both the gamble and the no gamble alternatives, that is, \$10. The expected utilities differ significantly. Respectively, the expected utilities of not gambling $EU_{NG}$ and $EU_G$ are as follows:

$$EU_{NG} = U(\$10)$$
$$EU_G = 0.5U(\$0) + 0.5U(\$20)$$

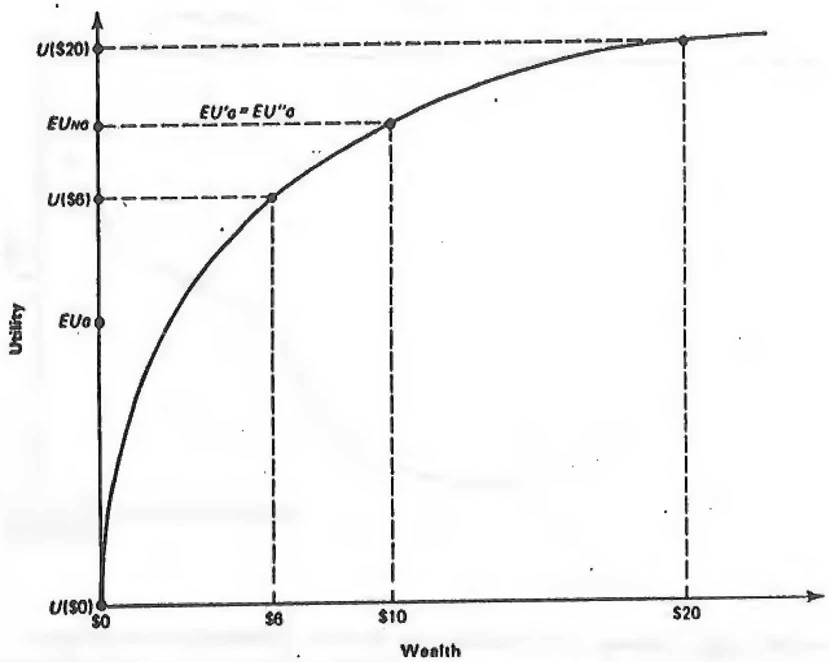



FIGURE 3-5

The respective positions are shown on the vertical axis, revealing that the individual will prefer not to gamble.

What will induce this risk-averse person to gamble? Two possibilities spring to mind: (1) keep the stakes the same, but change the odds in the gambler's favor or (2) keep the fifty-fifty odds, but change the stakes in the gambler's favor. For solution 1, consider the probability of

a Winning = 0.85
b Losing = 0.15

Therefore,

$$EU'_G = 0.15U(\$0) + 0.85U(\$20)$$

These odds have been engineered such that $EU_{NG} = EU'_G$, revealing that the individual is indifferent between not gambling and gambling at these favorable odds. If the odds are further improved in the gambler's favor, she would definitely prefer to gamble.

Much the same result can be achieved by changing the stakes instead of the odds, i.e., solution (2). Consider the gamble to involve a win of $10 with a probability of



0.5 and a loss of $4 with a probability of 0.5. The expected utility of gambling is now

$$EU''_G = 0.5U(\$6) + 0.5U(\$20)$$

which, again, has been engineered to give the same utility value as not gambling, as shown on the vertical axis of Figure 3-5.

In each of these cases, the expected value of the gamble has been increased in order to make the gamble more attractive relative to the no gamble choice. Expected wealth without gambling is $10. In contrast, the expected values, respectively, of (*1*) and (2) are

$$EV_{(1)} = 0.15(\$0) + 0.85(\$20) = \$17$$
$$EV_{(2)} = 0.5(\$6) + 0.5(\$20) = \$13$$

This increase in expected value is necessary in order to induce a risk averter to gamble, and as such, it is analogous to the risk premium discussed in connection with insurance.

### Resolution of the Insurance and Gambling Paradox

This analysis of gambling leaves us with a problem in explaining why people, in reality, choose to gamble. Even more puzzling is the fact that the same people often gamble and insure. The weekend visitor to Las Vegas or Atlantic City will return home and promptly pay the insurance premium for his or her home or life policy. Let us first address the simpler issue of why an individual might choose to gamble. The difficulty in explaining gambling is enhanced when we consider that most gambling is undertaken at unfair odds; after all, the bookmaker and the casino must come out ahead to stay in business.

1 One explanation of gambling is simply that gamblers are not risk-averse; their utility curves are convex rather than concave. For such individuals, gambling is rational behavior. However, we observe many people who gamble but otherwise appear to behave in a risk-averse manner. Some possible explanations of gambling that do not upset the notion of risk aversion are as follows:

2 Gamblers act on subjective rather than objective probabilities. Gamblers often believe they have a system or that they can read the cards or have a superstitious belief in "Lady Luck."

3 The attraction of gambling is not confined to the money outcomes; other forms of consumption also arise. A weekend in Las Vegas may be considered to be good entertainment, and any loss in gambling is the "price" of that entertainment. The unfair odds are comparable with the price of a theater ticket or the cost of a football game.

4 Another explanation, associated with the British economist G. L. S. Shackle, (1938) is that individuals do not respond mathematically to the odds, but instead focus their attention on particular outcomes that exert an undue influence on their decision.

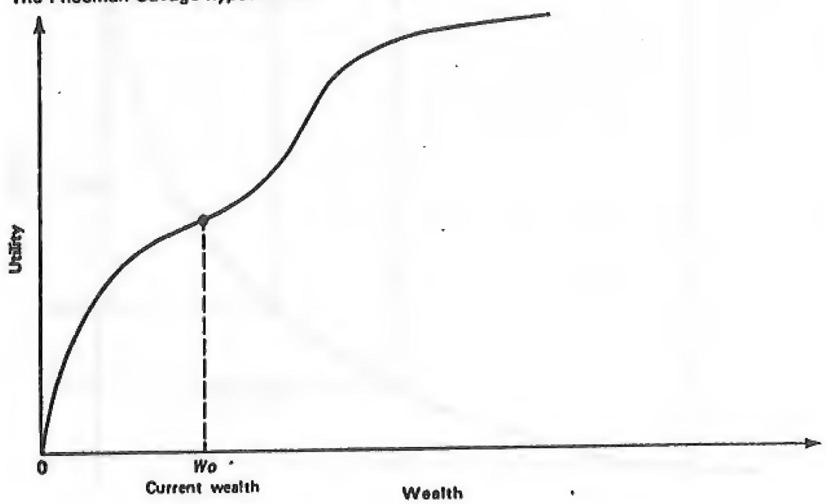
For example, with gambling, we might be blinded by the thought of winning; with insurance, we might be consumed by the image of our house reduced to ashes.

5 The previous explanation offered a possible resolution of why people might simultaneously gamble and insure. A more direct attack on this paradox was undertaken by Friedman and Savage (1948) and by Markowitz (1952), who suggested that the utility function may not be uniformly concave. A brief examination of the Friedman-Savage analysis follows.

Figure 3-6 shows a utility curve that is concave over most of its length, but it also exhibits a convex section over part of its range. It is important that the point at which the curve changes its curvature from concave to convex is current wealth $W_O$. Thus, for any reduction in wealth, the relevant section of the curve is concave (i.e., in the range $0–W_O$). Insurance relates to events that will reduce wealth from $W_O$, and since this range is concave, the individual responds in a rational risk-averse manner by purchasing insurance. For wealth increases (i.e., above $W_O$), the relevant section of the curve is convex. Since gambling offers us the prospect of wealth increases, this section of the utility curve is relevant in analyzing such decisions. The curve reveals risk preference in this range, and accordingly, it predicts that gambling will be undertaken at fair or even unfair odds. Subsequent debate on this issue has centered on whether the curve will eventually return to the concave form at very high wealth levels and on what form of gambling would appear to be rational from this model (high odds and high stakes or low odds and small stakes).

Some recent experimental evidence by Shoemaker and Hershey (1980) suggests

**FIGURE 3-6**
The Friedman-Savage hypothesis.






that the context in which a decision is presented may affect the decision. For example, consider the following choices:

$$A = \$100 \text{ with certainty} \qquad B = \begin{cases} \$120 & 0.8 \text{ probability} \\ \$0 & 0.2 \text{ probability} \end{cases}$$

The selection might represent the choice about purchasing insurance. Current wealth is \$120, but there is a 20% chance that it will be totally destroyed. For a premium of \$20, insurance can be purchased that gives the individual \$100 with certainty. The same prospects could be presented as a speculative business opportunity. The individual can invest \$100. There is an 80% chance that this will earn a return of \$20, giving the investor a final wealth of \$120, but there is a 20% chance that the investor will lose everything. Although these two stories are very different, the mathematical descriptions of the two problems are identical. Laboratory evidence reveals that subjects often respond differently when the context of the decision is changed.

The ideas discussed here reveal that actual decisions may be a little more complex and may have other dimensions than revealed by simple application of the expected-utility rule. Nevertheless, many of these ideas represent attempts to describe actual behavior, which may not always be rational or which may reflect unusual risk preferences, rather than attempts to identify rational optimizing behaviors. Our task is to prescribe a decision rule that will assist in identifying rational risk management decisions, and for this purpose, the expected-utility rule provides keen insights. Thus our main interest lies in the prescriptive value of the rule rather than its descriptive value.

### Partial Insurance

In the insurance decisions presented earlier, the choice was of an "all or nothing" nature. Insurance was purchased or it was not. More useful is a consideration of intermediate solutions. Various devices are currently available through which risk can be shared between insurer and insured. One common practice is to write an insurance policy with a deductible that assigns liability for the first *k* dollars of each loss to the insured, the insurer picking up the residual. Alternatively, the insurer may cover an agreed proportion of each loss, leaving the residual proportion to the insured. Other policies place an absolute limit on the insurer's liability. With such devices, the risk may be shared by the parties to the insurance contract, and the appropriate decision is one of degree. The question: How much insurance should be purchased? replaces the question: Whether insurance should be purchased?

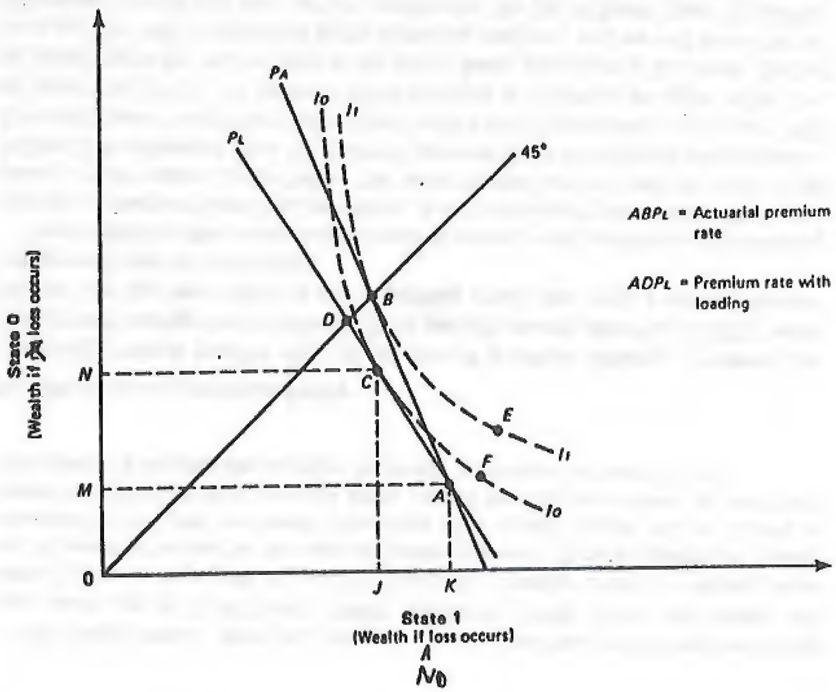
Consider here that the risk is divided on a proportionate basis. The individual has an initial wealth of \$120, of which \$20 is nonrisky and \$100 represents some asset that is subject to possible destruction with a probability of 0.25. Only total loss is assumed in this example. The reader will recognize this example (see Figure 3-3). In the earlier discussion, only full insurance was available, and it could be purchased at a premium of \$35, the expected value of loss being \$25. Let us view full insurance

as the purchase of \$100 of compensation in the event of (total) loss. However, we might have insured half the loss [i.e., purchased \$50 compensation in the event of (total) loss] or a quarter of the loss (purchased \$25 of compensation), etc. We may think of the degree of coverage as variable, the relevant issue being: How many dollars of protection should we purchase for the contingent \$100 loss?

To promote this example, the premium is scaled according to the number of dollars of protection purchased. If the premium is scaled according to the expected value of the insurer's claim payment, it will be 25 cents per dollar of compensation purchased, reflecting the 25% probability of loss. Full coverage would cost $\$100 \times 0.25 = \$25$; 50% coverage would cost $\$50 \times 0.25 = \$12.50$, etc. However, in the example given earlier, there was a premium of \$35 for full coverage, reflecting a premium rate of 35 cents per dollar of insurance purchased. What is the optimal level of insurance to be purchased?

This problem will be tackled using a rather different diagram. Figure 3-7 shows individual wealth in the event of a loss and in the event of no loss. Each possible event is called a *state of the world*. State 0 refers to loss, and state 1 refers to no loss. The respective axes show the respective wealth levels, and the initial contingent wealth position with no insurance is shown at position $A$. If the actuarially fair premium is charged, full insurance can be purchased for \$25. With such full insurance, the in-

**FIGURE 3-7**






dividual will have \$95 regardless of whether a loss occurs or not. This position is represented at position $B$. The 45° line shows equal wealth in both states and, as such, plots out full-insurance positions. Only by purchasing full insurance will wealth be independent of which state arises. However, the actual premium is 35 cents per dollar of insurance purchased. Thus starting at position $A$, each 35 cents of income sacrificed in state 1 in insurance premium will purchase \$1 of compensation in state 0 (loss payment from the insurer). The insurance possibilities are shown on the line $ADP_L$. Full insurance is shown at position $D$ (in which case, wealth would be \$85 with certainty), but the individual can select partial insurance along segment $AD$. Position $C$ represents a partial-insurance solution because it reveals greater wealth in the event of no loss than in the event of loss. In other words, compensation would be insufficient to provide full indemnity for the loss.

Now let us establish what level of insurance would be optimal at different premium rates. The solution of this problem requires some knowledge of the individual's risk preferences. These are represented by a set of curves such as $I_0$ and $I_1$, which are indifference curves. To understand these curves, consider two points such as $B$ and $E$. These are on the same curve, indicating that our subject is indifferent between the wealth combinations represented by these points. Notice the tradeoff involved in a move such as that from position $E$ to position $B$. Maintaining the same level of utility (i.e., indifference) requires that loss of income in the contingent state 1 be compensated by increased income in the contingent state 0. Despite such tradeoffs at a constant utility level, more wealth in both states is preferred to less wealth in both states; thus position $E$ is clearly preferred to position $F$. Therefore, position $F$ must offer a lower satisfaction level than position $E$, and consequently, it is on a lower indifference curve, that is, $I_0$. Summarizing, all positions on a given indifference curve represent the same level of utility, but any point on an indifference curve such as $I_1$ is preferred to any point on a lower indifference curve such as $I_0$. Indifference curves denote increased satisfaction as they move away from the origin. The convexity of the curves to the origin indicates aversion to risk. Thus a set of indifference curves can represent the same concepts as a utility function.

We know from the Bernoulli principle that a risk averter will choose to fully insure at an actuarially fair premium. Thus at premium rate $ABP_A$, the subject chooses position $B$ over partial-insurance solutions on the segment $AB$. The preference for position $B$ is shown by the fact that it is on the highest possible indifference curve that can be obtained given the opportunities along $AB$. The higher premium rate of 35 cents per dollar is shown by the line from $ADP_L$. However, the subject can no longer achieve the satisfaction level denoted by indifference curve $I_1$. The best that can be achieved is to just reach curve $I_0$, which offers lower satisfaction than $I_1$. This is the highest indifference curve that can be reached, as indicated by the tangency point $C$. Thus point $C$ is the optimal position. Notice that point $C$ is below the 45° line, indicating that less than full insurance is purchased. In fact, the insured pays $JK$ dollars in premiums to receive $MN$ dollars in compensation should the loss occur. Notice that the subject could have chosen not to insure (position $A$) or to fully insure at this premium rate (position $D$). However, both these positions are below indifference curve $I_0$, indicating lower utility than position $C$.

This diagrammatic approach illustrates an idea that was proved mathematically some years ago by both Mossin (1968) and Smith (1968): *that a risk averter will normally choose to partially insure if the insurance premium includes a loading factor that is positively related to the expected claim payment.* Since premiums are usually structured in this way, we would expect that most rational people will not choose to insure everything in sight, but will retain part of the risk themselves. As with most other goods, it appears that as price increases, we choose to purchase less.

### The Design of an Insurance Policy

The utility concept predicts that, in the face of premium loading, a rational risk-management strategy for an individual is to partially insure rather than to fully insure. We can use the same ideas to say something about how such a partial-insurance arrangement may be constructed.

As mentioned earlier, various devices are commonly used to structure a risk-sharing strategy between an insured and an insurer. A *deductible policy* covers only the surplus above a stated value. Losses below this value are not covered, and the value of the deductible is subtracted from the insurer's payment if the loss exceeds the deductible. A *proportionate coinsurance policy* simply pays a stated proportion of all losses. With an *upper-limit policy*, the insurer's claim payment is limited to an upper value; i.e., the insurer pays the value of loss or the policy limit, whichever is the lower. Each of these policies will be illustrated in the following example, and we will show that the risk-averse insured will have definite preferences for the different types of policies. The preferences are demonstrated in the example using the expected-utility rule. (A more general proof of these rankings can be derived mathematically, although this goes a little beyond the current text; see Arrow, 1963; Raviv, 1979; and Doherty, 1981.)

Consider an individual with an initial level of wealth of $200 (we can denominate this in thousands of dollars or millions of dollars if this is more appealing), which is tied up in physical assets, such as a home, furniture, etc. In addition, he or she has $60 in cash. The physical assets are subject to the possibility of damage or destruction, the appropriate information being given in the first two columns of Table 3-1. Thus, there is a

50% chance of no loss
10% chance of a loss of $20
20% chance of a loss of $40, etc.

The expected value of loss is $40. If the subject were fully insured, at a premium of $60, he or she would receive an insurance payment equal to the value of loss, as shown in the third column. The final wealth (i.e., after any loss has occurred and the insurance payment is made) is shown in the fourth column. Final wealth is calculated as the initial wealth ($260) minus the premium ($60) minus the loss plus an insurance payment. Since the subject is fully compensated, the final wealth does not vary from $200. However, our problem here is to choose between the different partial-insurance strategies shown in the remaining columns of the table. Respectively,





1 The *deductible policy* pays the value of loss minus $20, or $0, whichever is higher. Thus with a loss of $20, nothing is paid; with a loss of $40, the policy pays $20. The expected policy payment is $30, and the final wealth shown has an expected value of $205.

2 The *proportionate coinsurance policy* simply pays 75% of loss, irrespective of size. The insurance payments and final wealth are shown accordingly. Notice that this policy has been designed so that the expected value of the policy payment is again $30, and accordingly, the expected value of final wealth is again $205.

3 The *upper-limit policy* pays the value of loss or $100, whichever is lower. Again, the respective values of policy payment and final wealth are $30 and $205.

For each policy, the premium is $45, which is 150% of the expected policy payment. Which of these policies is preferred? On the face of it, they appear to give equal value for money, since the policies have the same actuarial value and the same premiums have been charged. (If the insurer had a well-diversified portfolio of such policies, then the same premium could be charged for each type, since long-run claim costs are identical.) The expected-utility rule will now be evoked to arrive at a ranking. To undertake this task, we will assume a particular form for the utility function that

**TABLE 3-1**
**EXAMPLES OF PARTIAL INSURANCE POLICIES**

| | | Full insurance | | $20 deductible | |
|---|---|---|---|---|---|
| Probability | Loss | Insurance payment | Final wealth | Insurance payment | Wealth |
| 0.5 | 0 | 0 | 200 | 0 | 215 |
| 0.1 | 20 | 20 | 200 | 0 | 195 |
| 0.2 | 40 | 40 | 200 | 20 | 195 |
| 0.1 | 100 | 100 | 200 | 80 | 195 |
| 0.1 | 200 | 200 | 200 | 180 | 195 |
| Expected value: | 40 | 40 | 200 | 30 | 205 |

| | | 75% coinsurance | | $45 upper limit | |
|---|---|---|---|---|---|
| Probability | Loss | Insurance payment | Wealth | Insurance payment | Wealth |
| 0.5 | 0 | 0 | 215 | 0 | 215 |
| 0.1 | 20 | 15 | 210 | 20 | 215 |
| 0.2 | 40 | 30 | 205 | 40 | 215 |
| 0.1 | 100 | 75 | 190 | 100 | 215 |
| 0.1 | 200 | 150 | 165 | 100 | 115 |
| Expected value: | 40 | 30 | 205 | 30 | 205 |

*Note:* Final wealth is calculated as: Initial wealth − Insurance premium − loss + policy payment

ensures that the shape concurs with the risk-averse, concave form illustrated earlier. The form chosen is[2]

$$U(X) = X^{0.8}$$

To calculate expected utility of wealth $EU$, we see that

$$EU = \sum_i P_i U(X_i)$$

The calculations are as follows:

1 Deductible:

$$EU_D = 0.5(215)^{0.8} + 0.1(195)^{0.8} + 0.2(195)^{0.8} + 0.1(195)^{0.8} + 0.1(195)^{0.8}$$
$$= 70.68$$

2 Coinsurance:

$$EU_C = 0.5(215)^{0.8} + 0.1(210)^{0.8} + 0.2(205)^{0.8} + 0.1(190)^{0.8} + 0.1(165)^{0.8}$$
$$\doteq 70.66$$

3 Upper-Limit:

$$EU_{UL} = 0.5(215)^{0.8} + 0.1(215)^{0.8} + 0.2(215)^{0.8} + 0.1(215)^{0.8} + 0.1(115)^{0.8}$$
$$= 70.55$$

A definite ranking appears. The deductible policy is preferred to a coinsurance policy having an equal expected policy payment. In turn, the coinsurance policy is preferred to an upper-limit policy having the same expected policy payment. This preference reveals that the deductible policy is a more effective instrument for containing the riskiness in the insured's final wealth. Certainly, the deductible exposes the insured to some risk, but an upper value is placed on his or her loss. However, the upper-limit policy limits the insurer's payment, leaving the insured's potential loss quite open-ended. Thus, if we compare the final-wealth columns under the respective policies, the final wealth under the deductible policy exhibits the lowest degree of risk and the final wealth under the upper-limit policy is the most risky. A cautionary note is given: The expected utilities are rather close in value. It should not be concluded that differences in utility are marginal. The expected-utility rule is a ranking device only.

[2]The reader may wish to experiment with other types of functions that also exhibit the concave property, for example,

$$U(X) = X^{0.5} \quad \text{or} \quad U(X) = \log X$$



This example may also be used to illustrate that partial coverage is often preferred to full coverage. In the previous section we demonstrated that if the premium loading is positively related to the expected value of claim payment, some degree of partial coverage is optimal. We have not tried to ascertain what level of risk sharing is optimal; thus the $20 deductible may be inferior to a $30 deductible. However, it is of interest to see whether the various forms of partial coverage are preferred to the full-insurance policy. The respective premiums for each of the policies are set at 150% of the expected policy payment; thus in each case the loading is 50% of the expected payment. This satisfies our premium criterion, since the loading is positively related to the expected payout. The utility index for full coverage is established easily, since final wealth is $200 with certainty. Expected utility is

$$EU_{FC} = 200^{0.8} = 69.31$$

Clearly, each of the partial-insurance strategies is preferred to full insurance. This is quite compatible with our previous result, although it is still possible that some other levels of partial insurance may prove even better.

## ALTERNATIVE DECISION RULES

### Problems with the Expected-Utility Rule

The expected-utility rule is a very useful device for helping to condition our thinking about risky decisions, because it focuses attention on the types of tradeoffs that have to be made. Furthermore, the results it generates are useful as first approximations in the search for risk-management solutions. Certainly, it makes sense to transfer risk if there is no sacrifice in terms of expected return. When the transfer of risk is costly in terms of expected return, somewhat lower insurance purchases appear to be justified. And so on! However, the expected-utility rule and the literature it has spawned do have several limitations that reduce their relevance for risk-management purposes:

1 To calculate expected utility, we need to know the precise form and shape of the individual's utility function. Typically, we do not have such information. Usually, the best we can hope for is to identify a general feature, such as risk aversion, and to use the rule to identify broad types of choices that might be appropriate.

2 The rule cannot be applied separately to each of several sets of risky choices facing an individual. For example, consider three choices:

*Choice 1:* Purchase automobile insurance *or* do not purchase automobile insurance.

*Choice 2:* Purchase home insurance *or* do not purchase home insurance.

*Choice 3:* Purchase risky shares of stock *or* put money in riskless bank account.

Decisions such as these are interdependent, as we shall see in the bulk of this book. However, the expected-utility rule has normally been applied to each choice as though it were independent of others.

3 The third problem is really a composite of the first two. For corporate risk management, it may not be possible to consider a utility function for a firm as though the firm were an individual. Much of risk management has been developed on the simplistic, and possible erroneous, assumption that a firm can be treated as a risk-averse individual. However, a firm is a coalition of interest groups, each having claims on the firm: shareholders, bondholders, managers, employees, customers, agents, etc. The decision process must reflect the mechanisms with which these claims are resolved and how this resolution affects the value of the firm. Furthermore, the risk-management costs facing a firm may be only one of a number of risky prospects affecting the firm's owners and other claimholders. The expected-utility rule is not an efficient mechanism for modeling the interdependence of these sources of risk.

We will briefly outline alternative decision rules that can be used for risky choices. The first of these rules, the *mean variance rule*, will occupy a central role in much of this book. The second rule, *stochastic dominance*, is more sophisticated, but is presented in sketchy form only because it is more difficult to apply. Some readers may choose to skip the stochastic dominance section, and this will not destroy continuity. Stochastic dominance uses cumulative probability distributions, so the reader who is unfamiliar with this concept may wish to defer or omit this treatment.

## The Mean-Variance Rule

The *mean* in the name given to this rule relates to expected value. Expected value already has been used as a simple decision rule. Its drawback is that it ignores the riskiness of the different alternatives, as dramatically illustrated by the St. Petersburg paradox. Instead of developing the subtle and indirect treatment of risk implied by expected utility, we could instead find a direct measure of risk and then make our choice on the basis of *both* expected value and risk.

Consider the alternatives shown in Table 3-2. Assume that people are risk-averse, but prefer a higher expected value to a lower expected value. Under such conditions, which probably apply to most of us, clearly choice A is better than choices B, C, or D; choice C is better than choice B; and choice D is better than choice B. In each of these rankings, the preferred alternative scores higher on at least one of the two criteria, but it does not score lower on either criteria. For example, choice A is preferred to choice D on expected value, but they score equally on risk. Notice, however, that the risk/expected-value rule does not rank all alternatives. Choices C and D cannot be ranked. Choice D is preferred to choice C on the risk criterion, but on the basis of expected value, choice C has a better result. Thus the risk/expected-value rule will not always work.

**TABLE 3-2**

| Choice | Risk | Expected value |
|---|---|---|
| A | Low | High |
| B | High | Low |
| C | High | High |
| D | Low | Low |



In order to make this approach operational, we need to be able to measure risk. From reading introductory statistics textbooks, the reader will recall that various measures of risk are available. Consider a risky prospect such as

$$A = \begin{cases} \$10 & {}^1/_3 \text{ probability} \\ \$20 & {}^1/_3 \text{ probability} \\ \$30 & {}^1/_3 \text{ probability} \end{cases}$$

The risk is implicit in the spread of possible values. One easy measure is simply the *range* of values, i.e., from \$10 to \$30. The range is the highest possible value minus the lowest possible value. Clearly, this gives some idea of spread, but it ignores probabilities. Among the various alternative contenders for a risk measure, the one with the most convenient properties is standard deviation (or its square, variance). Standard deviation $\sigma$ is defined as

$$\sigma = \left[\sum_i P_i(X_i - \bar{X})^2\right]^{1/2}$$

where $P_i$ = probability of outcome $X_i$

$\bar{X}$ = expected value, which equals $\sum_i P_i X_i$

The expected value $EV$ of prospect $A$ is

$$EV_A = {}^1/_3(\$10) + {}^1/_3(\$20) + {}^1/_3(\$30) = \$20$$

and the standard deviation is

$$\sigma_A = [{}^1/_3(\$10 - \$20)^2 + {}^1/_3(\$20 - \$20)^2 + {}^1/_3(\$30 - \$20)^2]^{1/2} = 8.165$$

The variance is

$$\sigma_A^2 = 8.165^2 = 66.67$$

This measure will be used extensively in later chapters. For the moment, we will simply examine one or two general properties of the mean-variance rule. Consider the choices outlined in Table 3-2. We can represent each of these as a probability distribution, showing the range of possible values on the horizontal axis and the associated probabilities on the vertical axis. The four choices are represented in this way in Figure 3-8. Choices A and D have the same shape or spread, indicating a low level of risk;

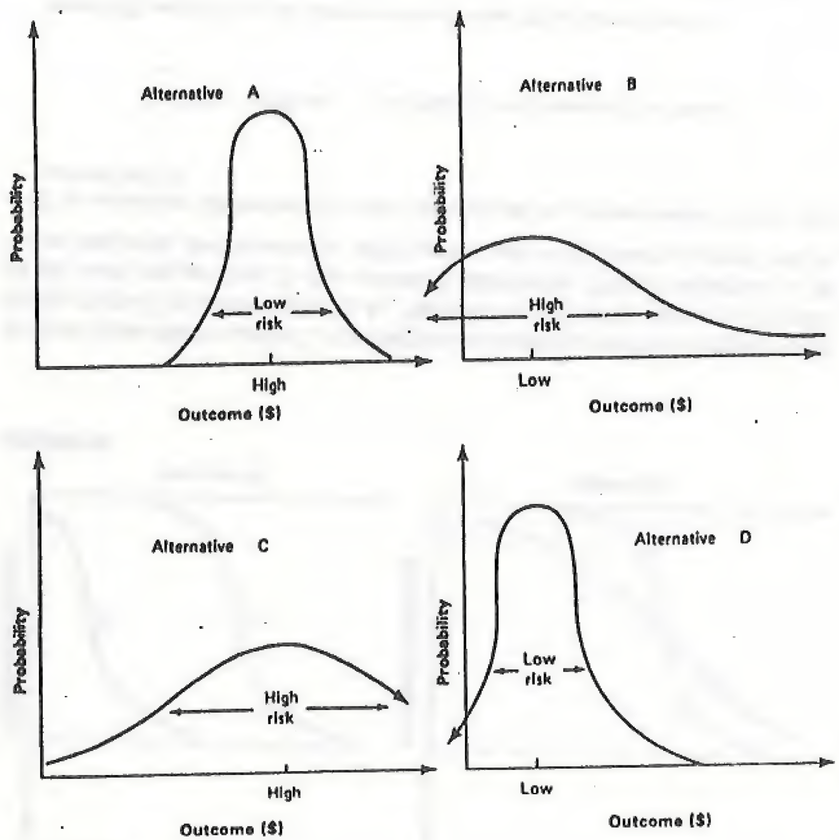



**FIGURE 3-8**

in each case, the possible values are closely clustered around a central value. The central, or expected, values of the two distributions are different. The expected value of choice A is high, whereas the expected value of choice B is low. Alternatives B and C have the same shape, the wide spread indicating high risk. However, whereas choice B is centered on a low expected value, choice C centers on a high expected value. Comparison of different pairs of these alternatives will confirm the rankings we derived earlier.

Both the expected-utility rule and the mean-variance rule require the decision maker to be able to estimate the possible outcomes and their respective probabilities. Such measures are feasible in principle and often in practice. However, this is the only input required for the mean-variance rule, in contrast with the expected-utility rule, which also requires specification of the utility functions. This operational advantage is achieved at some cost, since we find that not all risky choices can be ranked. We



derive only a limited ranking that applies to all risk-averse individuals. This rule has proved very useful in analyzing investor behavior and, in consequence, in valuing the firm.

### Stochastic Dominance

Like the mean-variance rule, stochastic dominance describes a decision procedure that is applicable for risk averters and does not require specification of the individual utility function. As with mean variance, it can rank some, but not all, choices. However, it has one particular advantage over mean variance. Expected value and standard deviation or variance are arbitrary values for the central tendency and riskiness of a prospect. Furthermore, they describe only certain features of a prospect. Often a risky prospect cannot be represented by a smooth or symmetric curve, such as those shown in Figure 3-8, but rather the curve will be irregular and asymmetric. Such features may be relevant to an individual's choice, but they are ignored by the mean-variance rule. Stochastic dominance is designed to overcome such problems, although it is more difficult to use and may sometimes fail to produce clear rankings, even where mean variance succeeds.

Stochastic dominance is really a set of decision rules that applies to progressively more restrictive groups. A full mathematical specification is beyond our scope. However, an intuitive understanding is easily presented.

Consider distributions A and D shown earlier in Figure 3-8. These are represented in Figure 3-9 on the same diagram, and immediately below, the respective cumulative distributions are presented. The cumulative probability distribution shows the probability that any given outcome will be equal to or less than a given value. Examples of probability and cumulative probability distributions are developed in Chapter 4, and the reader who is unfamiliar with these distributions may wish to defer discussion. The preference of A over D is pretty clear. Both have the same shape, indicating the same level of risk, but A clearly centers around a higher expected outcome. The cumulative probability distributions reveal that distribution A is always equal to or below distribution D. In essence, prospect D exhausts "probabilities" at low levels of wealth. In prospect A, the probability is allocated at higher wealth levels; thus the distribution appears to be shifted to the right. This reveals the *first rule for stochastic dominance:*

> If the cumulative distribution of A is equal to or below that for D for every level of wealth, then prospect A dominates (is preferred to) prospect D.

This rule is not restricted to those who are averse to risk, but it does apply to all who prefer greater wealth to lesser wealth. This probably describes us all, and the rule is quite general.

Now look at the right-hand side of Figure 3-9. Distributions for A and C are extracted from Figure 3-8 and are presented on a single graph. The distributions have the same expected value, but choice C exhibits considerably more risk. Choice A should be preferred to choice C. The appropriate stochastic dominance rule is developed from the cumulative distribution. It will be seen that the cumulative distribution for choice

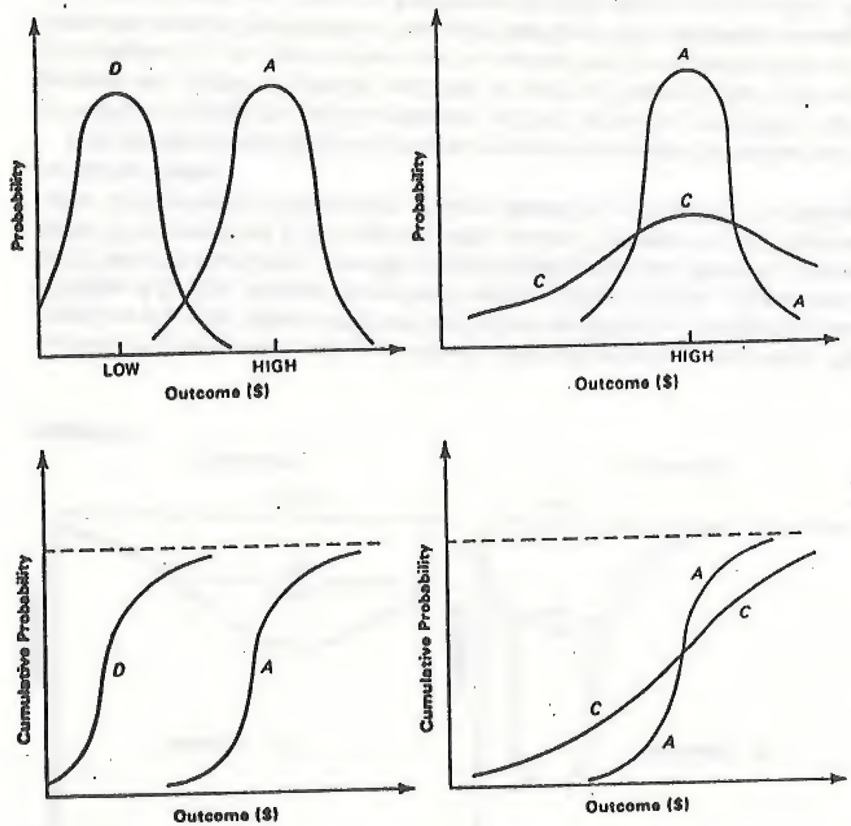



FIGURE 3-9

A starts below that for choice C, the distributions intercept, and then at higher wealth levels, choice C is lower than choice A. The lower level of risk in choice A is revealed by the much steeper slope in the cumulative distribution. This is anticipated in the *second stochastic dominance rule*, which applies only to risk-averse decision makers:

> If the cumulative distributions of A and C intercept one or a greater number of times, A is preferred to C if
>
> $$\int_{-\infty}^{x} [C(X) - A(X)]\, dx \qquad x \geqslant 0 \text{ for all } x \text{ with inequality for some } x$$
>
> where $C(X)$ and $A(X)$ are the cumulative distributions for prospects $C$ and $A$

The application of this rule is a little complex, although in Figure 3-10 we give an



example of how it might apply to a simple risk management problem. If full insurance is purchased, the resulting distribution of wealth for an individual will exhibit no risk; wealth level $W_0$ will be achieved with certainty. An appropriate cumulative distribution is shown as $I$ in Figure 3-10. If no insurance is purchased, wealth will be risky, and an appropriate distribution is shown as $N$. The distributions intersect just once. Under such circumstances, second-degree stochastic dominance will apply if the premium for insurance is actuarially fair. The full-insurance distribution is preferred to the no-insurance distribution. This is compatible with the Bernoulli principle.

Our treatment of stochastic dominance is sketchy and will not be developed further. The interested reader might consult the literature review provided by Bawa (1982) for further applications, including insurance applications.

## PROBLEMS IN INSURANCE MARKETS

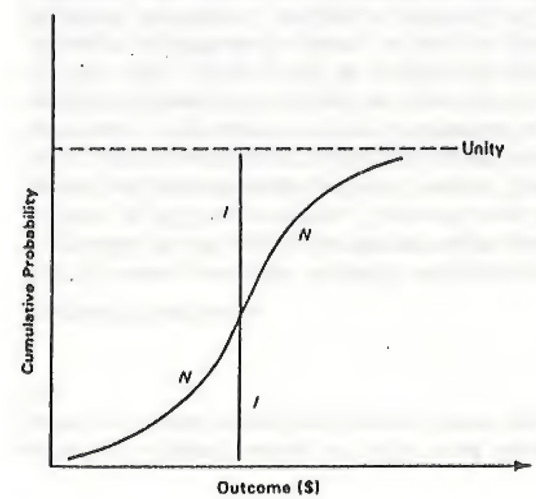
### Moral Hazard

Economic analyses have been used to address the risky choices that often face individuals, particularly the risk management choices that are of interest here. However, similar analyses have been used to enrich our understanding of how insurance markets function. Two unusual and possibly disturbing issues are examined: moral hazard and adverse selection.

The term *moral hazard* is emotionally loaded. This is quite unfortunate, since many activities that fall under such a label are neither improper nor illegal. Any ethical connotations placed on this term will serve only as a barrier that clouds our understanding of the real issues.

*Moral hazard* relates to the behavioral effect that insurance might have on the level

FIGURE 3-10



of effort or expenditure devoted to loss-reducing activities. If an individual or a corporation owns property, the ultimate value of that ownership claim depends on the value of the property. The property value might be reduced by the risky events we consider here, such as fires, explosions, etc. The value of individual or corporate wealth is similarly affected by claims by employees, agents, and third parties for defective products, industrial injuries, and other torts and statutory liabilities. Any expenditure allocated to reduce the probability of such loss, or the size of any loss that might occur, will benefit the owner and increase the value of his or her ownership rights. Thus ownership of a sprinklered building is more valuable than ownership of an otherwise identical, unsprinklered building.

The owner has an incentive to protect his or her ownership rights, and it makes sense to undertake those forms of protection which yield benefits in excess of costs. To understand moral hazard, it is useful to see how this incentive is affected by the purchase of insurance:

1 The insurance policy transfers the cost of losses from the owner to the insurer. Any reduction in this cost will accrue to the insurer, who will thereby derive benefit from loss-reduction efforts. With full insurance, the owner will be indifferent as to whether the loss occurs or not (unless, of course, there are nonfinancial dimensions to the loss that cannot be compensated, such as pain and suffering and sentimental value). For the owner, the payoff to any expenditure on loss reduction is reduced possibly to zero by the insurance policy. Accordingly, there is a reduced or zero incentive for the owner to protect the property or other rights that are insured. The transfer of the cost of loss under an insurance policy reduces or removes incentives for loss reduction.

2 This effect may be modified by the pricing structure or the contractual conditions of the insurance policy. If the insurer is able to monitor any loss-prevention efforts, it may structure the premium to restore loss-reduction incentives. Premium reductions may be allowed for installation of preventative devices and better safety practices, and premium increases may be included for adverse features. For example, fire insurance schedules usually specify premium adjustment for sprinklers, fire alarms, construction standards, storage of hazardous materials, etc. Even if the insurer cannot directly observe safety practices, it may indirectly reward good practice by calculating future premiums with respect to past loss records. Experience rating and retroactive rating have this effect.

3 Much the same effects can be achieved by imposing conditions in the insurance policy that make insurance coverage conditional on certain safety practices. Possibly this practice is less flexible than direct premium adjustment. The incorporation of both policy conditions and premium adjustments will serve to restore incentives for loss reduction that had been removed by the transfer of loss costs under the insurance policy.

The problem of moral hazard arises from an inequality in the information on loss reduction available to the insured and insurer. An insurer may be able to observe whether a building is sprinklered and the combustibility of the materials used in its construction, but it may not be possible to observe the standards of housekeeping and everyday work practices that may give rise to fire, explosion, and other hazards. Those practices which cannot be observed cannot be controlled by the insurer by pricing incentives or by policy conditions. The essential problem for the insurer is that it may not be able to charge the correct premium. As a consequence, insurers may be reluctant to offer insurance protection when they suspect moral hazard will be a problem.



Moral hazard has other dimensions, however. From the risk manager's viewpoint, moral hazard is simply rational economic behavior. This is not to say that risk managers should be sloppy about safety practices or try to deceive the insurer about loss expectancy. Rather it is simply an assertion that the rational risk manager would take account of any premium incentives in constructing a loss-prevention program. For example, in deciding on the type and form of fire prevention for a building, the premium savings offered by an insurer may represent a major part of the expected benefit to the insured. If the insurer does not offer substantial premium reductions, the effect may be to kill the fire safety program, even though such a program might well have been undertaken without insurance.

Much attention has been given in insurance literature to the adverse effects that insurance may have on the allocation of resources for loss reduction and preventative activities. The role of premium incentives has received relatively little attention. If premiums are sensitive to changes in the expected value of loss, they will serve the valuable purpose of conveying information to the insured on the effectiveness of different forms of loss reduction. Insurers essentially serve as a clearing house for loss-reduction information. By classifying loss data according to the identifiable features of different risks, insurers may compile information on the contributions of those features to the expected value of loss. This information may then be reflected in premium differentials, thereby providing information to the insured on the effectiveness of different forms of loss prevention. Several features of insurance market lend support to the view that insurers disseminate appropriate safety information:

1 Competition in insurance markets tends to lead to premium structures that adequately discriminate among risks according to loss expectancies.

2 Insurers often pool information to refine their data base.

3 Regulation of insurance markets usually aims toward supporting the concept of fair discrimination.

4 Insurers typically employ technical experts to survey risks in order to convey information to underwriters and to insureds.

If markets are reasonably efficient in disseminating loss-prevention information, then it would seem prudent for risk managers to look to the conditions of insurance protection to guide them in loss-prevention decisions. What may be labeled *moral hazard* may sometimes lead to an efficient allocation of loss-reduction resources.

### Adverse Selection

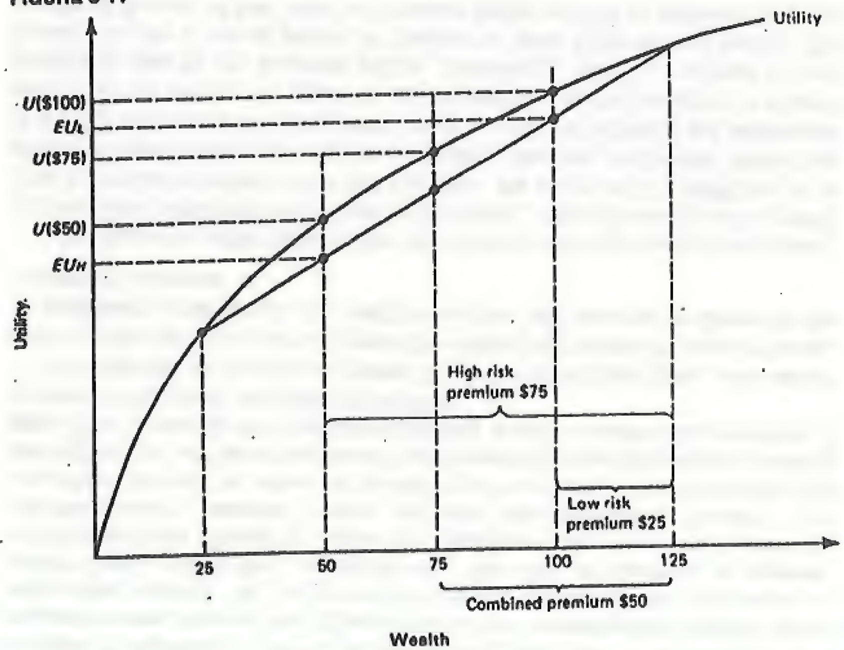
Like moral hazard, *adverse selection* may have the effect of reducing the supply of insurance services. Furthermore, it may prevent the insurance market from reaching a set of stable equilibrium prices. Adverse selection plays an important role in the

functioning of insurance markets, and an introductory understanding is useful for risk managers.

Both adverse selection and moral hazard arise from information asymmetry between the insurer and insured. Consider an insurer that writes a portfolio of policies. Insureds may be grouped into two classes: those with a high expected value of loss and those with low loss expectancy. If the insurer can distinguish insureds according to their respective loss characteristics, each can be charged a premium that reflects his or her expected value of loss. Thus insurers may use observable characteristics, such as automobile type, location, or age, to distinguish different risk groups of automobile policies. However, the insurer cannot always separate risk groups and will be forced, by lack of information, to charge the same premium. This results in an effective subsidy from low-risk insureds to high-risk insureds. The effects of this subsidy may destabilize the insurance market and reduce underwriting capacity, as shown below.

In Figure 3-11 we consider a portfolio of two groups of insureds: high risk and low risk. Each insured starts with a wealth level of $125, but a loss (total loss only) may reduce the wealth to $25. The groups differ in the probability of such a loss. For the high-risk group, the probability of loss is 0.75, resulting in an expected value of loss of $75. For the low-risk group, the probability loss is 0.25, resulting in an expected

**FIGURE 3-11**






loss of $25. If the insurer can distinguish between the two groups, respective competitive premiums of $75 and $25 may be charged (ignoring transaction costs). With premiums set at the expected value of loss for each insured, the Bernoulli principle asserts that each would fully insure. Thus, for the low-risk group, the utility of insuring and having wealth of $100 with certainty, that is, $U(\$100)$, is higher than the expected utility of not insuring $EU_L$. Thus,

$$U(\$100) > EU_L = 0.75U(\$125) + 0.25U(\$25)$$

and for the high-risk group,

$$U(\$50) > EU_H = 0.25U(\$125) + 0.75U(\$25)$$

The respective positions are shown on the vertical axis of Figure 3-11.

Now suppose that the insurer is unable to distinguish between high- and low-risk insureds. If there are equal numbers in each group, the breakeven premium will be $50. However, at this premium we see that the low-risk group will not insure because the utility of not insuring $EU_L$ is greater than the utility of insuring and having a wealth level of $75 for certain, that is,

$$EL_L > U(\$75)$$

Conversely, the high-risk group will find insurance to be a bargain and will rationally choose to insure, that is,

$$EU_H < U(\$75)$$

Thus the portfolio composition will change as low-risk insureds cancel their policies, leaving a portfolio of high-risk insureds (each having an expected cost of $75) and an inadequate premium.

In practice, the process will be somewhat smoother. There may be several risk groups, and coverage may be arranged on a partial basis. The insurer that averages premiums over a number of risk groups will find that it tends to lose the good risks as they cancel or reduce their coverage. The resulting change in the composition of the portfolio will cause the average premium to be inadequate, forcing the insurer to raise the premium. This serves merely to aggravate the flight of low-risk insureds, and so the process continues until only bad risks are left. This process is usually attributed to information deficiency on the part of the insurer, but the same effect may result from regulation designed to prevent insurers from using classification variables that are politically sensitive, such as sex or race. Indeed, one of the rare pieces of empirical evidence documenting this process relates to regulation-induced adverse selection (see Dahlby, 1983).

Competition between insurers may help reduce problems of adverse selection. Information on loss expectancies of individual insureds is of economic value to an

insurer. Armed with such information, an insurer can selectively attract low-risk insureds from a rival insurer that is unable to discriminate simply by offering a lower price and admitting only low-risk insureds. Thus competition will induce insurers to seek and compile information that will enable them to operate premium structures that adequately discriminate between risk groups. Of course, information will never be perfect, and adverse selection will never disappear. But in an actively competitive market, adverse selection will be reduced to a level that is compatible with the costs of information.

## SUMMARY AND CONCLUSION

The task of this chapter has been to present decision rules that can be used to rank competing risky choices. From these rules we have derived a set of preliminary risk-management strategies for individuals that will help focus our thoughts when we tackle the more complex corporate risk-management decisions.

The simplest decision rule considered is to choose that alternative which has the highest expected value. There is some long-run virtue in this rule. Its repeated use in many separate decisions will ultimately lead to wealth maximization. However, the rule fails to account for differences in the riskiness of competing prospects, as dramatically illustrated by Bernoulli's St. Petersburg paradox.

The expected-utility rule requires that the decision maker select that prospect with the highest expected utility. Under this rule, the possible outcomes are weighted according to their respective probabilities and according to the utility scale of the decision maker. The substitution of outcomes measured in utility terms for money outcomes ensures that individual risk preferences can be impressed on the decision process. This rule has been widely used in the economics literature, and it is used here to compare some simple risk-management choices.

The expected-utility rule can be used in conjunction with the assumption of aversion to risk to reveal a propensity to insure at fair prices (premiums are equal to the expected value of loss) or even at unfair prices. There is an equivalent aversion to gambling. This is not to assert that people will never gamble and will always insure, for both activities can be priced to encourage or deter demand. Under realistic pricing assumptions, partial-insurance solutions often appear more attractive than "all or nothing" attitudes toward insurance purchase. For insureds, deductibles appear to provide a more attractive mechanism for risk sharing than other devices, such as coinsurance or policy limits.

Despite its generality, the expected-utility rule does have the serious shortcoming that it requires the decision maker to specify his or her utility function. Thus while the rule establishes generalized patterns of behavior that are useful in guiding individual decisions, it lacks operational value for corporate risk-management purposes. An alternative is to use mean and variance to characterize risky prospects. This approach is introduced and will be used extensively in later chapters.

To set the background for later analysis, the economic concepts introduced in this chapter were used to illustrate two important issues that arise in insurance markets and which constrain the supply of insurance services. *Moral hazard* refers to the adverse incentives that insurance may convey to the insured to reduce the prospect of future loss. The problem may be severe if the insurer is unable to monitor the insured's behavior. Similar information problems may lead to *adverse selection*. If the insurer is unable to correctly discriminate between insureds in its pricing structure, it will effectively subsidize the higher risks. This will lead to disproportionately high demand from high-risk groups, causing a deterioration in the risk composition of the insurer's portfolio and escalating loss experience. The prospect of adverse selection again acts as a constraint on insurance supply or capacity.

We will turn our attention in the ensuing chapters to the identification and measurement of those risks which form the subject matter of risk management.



## QUESTIONS

1 Consider the following game. Peter is to turn over one card from a deck. If the card is a face card (ace, jack, queen, or king), Peter will win $100. If any other card is selected, Peter will win nothing. To play this game, Peter must pay an entry price of $25.
   a What is the expected value of the game
   b If Peter's utility function is

$$U = W^{0.5}$$

   where $U$ = utility and $W$ = wealth, should Peter enter the game? Assume Peter's starting level of wealth is $200.

2 Construct utility diagrams to display the following:
   a A risk averter will insure when faced with an actuarially fair insurance premium.
   b The risk premium (the difference between the insurance premium that a risk averse person is willing to pay and the expected value of the loss).
   c A risk preferrer will choose to gamble when the expected value of the gamble is equal to the entry price.
   d A person who may rationally choose to gamble and insure when he knows all probabilities.

3 Donne faces the following loss distribution:

| Loss | Probability |
|---|---|
| $ 250 | 0.4 |
| 500 | 0.4 |
| 1,000 | 0.2 |

An insurance policy can be purchased that
a has a $250 deductible for a premium of $300.
b covers 50% of all losses for a premium of $350.
c fully covers all losses for a premium of $550.
Donne's initial wealth is $2000 and his utility can be described by the function

$$U = \log W$$

where $W$ is final wealth. Which, if any, policy should Donne purchase?

## REFERENCES


Arrow, K. J.: "Uncertainty and the Welfare Economics of Medical Care," *American Economic Review*, vol. 53, 1963, pp. 941–973.

Bawa, V. J.: "Stochastic Dominance: A Research Bibliography," *Management Science*, vol. 28, 1982, pp. 698–712.

Borch, K.: *The Economics of Uncertainty*, Princeton Studies in Mathematical Economics, Princeton University Press, New Jersey, 1968.

Dahlby, G. G.: "Adverse Selection and Statistical Discrimination," *Journal of Public Economics*, vol. 20, 1983, pp. 121–130.

Doherty, N. A.: *Insurance Pricing and Loss Prevention*, D. C. Heath, Lexington, Mass., 1976.

———: "Stochastic Ordering of Risk Insurance Exchanges," *Scandinavian Actuarial Journal*, no. 4, 1981, pp. 302–308.

Friedman, M., and L. J. Savage: "The Utility Analysis of Choices Involving Risk," *Journal of Political Economy*, vol. 56, 1948, pp. 279–304.

Hirshleifer, J., and J. G. Riley: "The Analytics of Uncertainty and Information—An Expository Survey," *Journal of Economic Literature*, vol. 17, 1979, pp. 1375–1421.

Markowitz, H.: "The Utility of Wealth," *Journal of Political Economy*, vol. 60, 1952, pp. 151–158.

Marshall, J. M.: "Moral Hazard," *American Economic Review*, vol. 66, 1976, pp. 880–890.

Mossin, J.: "Aspects of Rational Insurance Purchasing," *Journal of Political Economy*, vol. 76, 1968, pp. 553–568.

von Neumann, J., and O. Morgenstern: *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, New Jersey, 1944.

Raviv, A.: "The Design of an Optimal Insurance Policy," *American Economic Review*, vol. 69, 1979, pp. 223–239.

Shackle, G. L. S.: *Expectations Investments and Income*, Oxford University Press, London, 1938.

Shoemaker, P. J. H.: *Experiments on Decisions Under Risk*, Kluwer Nijhoff Publishing, Amsterdam, 1982.

———, and J. C. Hershey: "Risk Taking and Problem Context in the Domain of Losses: An Expected Utility Analysis," *Journal of Risk and Insurance*, vol. 47, 1980, p. 111.

Smith, V.: "Optimal Insurance Coverage," *Journal of Political Economy*, vol. 76, 1968, pp. 68–77.

Zeckhauser, R.: "Medical Insurance—A Case Study of the Tradeoff between Risk Spreading and Appropriate Incentive," *Journal of Economic Theory*, vol. 2, 1970, pp. 10–26.

# CHAPTER 12

# DIVERSIFICATION AND INSURANCE

Chapter 5 examined why individuals might wish to purchase insurance. People are exposed to risk of financial loss. Such losses may arise if they crash their cars, if their houses burn down, if they lose their jobs, if they lose their wallets, if they are burgled or if they are sued for negligent behaviour. The financial uncertainty from such events is undesirable; people are usually "risk averse". Consequently, they are often willing to substitute the certain payment of an insurance premium for an uncertain financial loss. In this description, the risk facing the individual is unacceptable and is sold to the insurance firm for an appropriate premium.

Now look at the transaction from the viewpoint of the insurance firm. The insurance firm has chosen to accept risk. It receives a certain premium and faces a very risky prospect of paying for any future losses that may arise. Moreover, the insurance firm has accepted such risks from a large number of policyholders. Why is it that the risk that is unacceptable to the individual is perfectly acceptable to the insurance firm? Why is it that, in selling a large number of policies, the insurer is not swamped with risk?

This chapter will show how the insurance firm controls and diversifies risk. When risky cash flows are combined together in a portfolio, the risks do not "add up"; the total risk is less than the sum of the parts. Under ideal circumstances, the insurer may be able to predict with a high degree of reliability, how much it must pay in total claims to its policyholders. In these circumstances, the risk has all but disappeared for the insurer. This process is known as DIVERSIFICATION or RISK POOLING. Just how diversification works is the subject of this chapter.

## COMMON EXAMPLES OF DIVERSIFICATION

Examples of diversification arise in many aspects of life and we probably take most for granted.

(a) Consider alternative college grading methods. In most courses, students are assessed, not with a single examination, but with a combination of exams, tests, assignments and the like. If there were a single examination, would this truly reflect the ability of the student? Without intending to settle the score, critics of the single exam, often suggest that is not a good testing method since the student may have a unusually good (or bad) day, or the limited number of questions does not test the breadth of knowledge of the student. If there are many tests and exams, the argument goes, the student will have some days in which he does not perform up to his ability and some inspired days. These will tend to cancel out, or average out, and leave a final grade that is representative of ability. Averaging over the semester, it is claimed, is a truer test of ability.

From the student's viewpoint, the single exam is more risky since the grade reflects the "performance on the day". This risk is

reduced by spreading the grade over various tests and assignments. The student does not have "all his eggs in one basket".

This argument is not intended to settle the score against the single examination. Proponents often claim that it helps test the student's ability to integrate all the material from the course. Further, life itself is full of "one shot" chances and the single exam simply reflects this.

(b) In much the same vein, consider the difference between sports competitions that are organized on a "knockout" basis and those organized on a "league" basis. To focus more closely, consider the Stanley cup playoffs in ice hockey. The final two teams play up to seven games. If a team wins four games, it wins the cup. Why seven games, why not one game?

If the cup were decided one one single game between the final teams (presumably at some neutral location) the losing team could argue they were simply unlucky; they happened to have a bad day and all the "breaks" went against them. Though they would not vocalize such thoughts, the winners might quietly contemplate that fortune was with them; everything just happened to come out right on the day.

Spreading the issue over seven games helps to diminish the effects of pure luck on the outcome. Over seven games, each team may have the odd lucky win, or unlucky loss, but the underlying ability of the teams should show through in the end. The seven game series has spread the risk.

(c) An investor often will hold a portfolio of assets rather than placing all of his capital in a single stock/bond/painting/etc. If all capital is invested in a single asset, the investor is vunerable to movements in its price. For example, many stock prices are quite volatile. If the stock price happens ro rise after it is purchased, the investor will secure a capital gain which is fine and dandy. But if the price happens to fall, the investor will suffer a loss of capital. Given the volatility of stock prices, the capital gain or loss could be quite large. Such an investment is very risky.

Now, if the investor spreads his capital over a number of financial assets, e.g. a number of stocks, the risk will be reduced. Some stocks may bring capital gains since their prices rise. But it is probable that others will bring capital losses through falling prices. Over a large portfolio, these gains and losses may tend to offset each other thus stabilising the investment performance. Holding a portfolio avoids the risk of unusual price movements from holding single assets. The portfolio approach spreads the risk. Thus as the portfolio includes more and more securities, the actual rate of return achieved by the investor should be closes and closer to its expected value (i.e. there would be smaller and smaller risk of abnormally good or abnormally bad investment results).

This idea was introduced in Chapter 9 when considering the pricing of capital assets. However, there it was suggested that such diversification

works best when asset prices tend to move independently of each other. If all asset prices rose or fell together, then it is clear that the portfolio investment strategy would not benefit by offsetting capital losses on some stocks with capital gains on others. Common movement of stock prices implies that all stocks in the portfolio would bring capital gains or all would bring capital losses. Thus the success of diversification depends on the COVARIANCE, or common movement, between the prices of stocks in the portfolio.

(d) It is often useful to know certain characteristics of a population such as political attitudes, economic or physical characteristics. Suppose, for example, that we wished to know the average height of American females. We could only truely know the answer by measuring the height of every American female and dividing by the number involved. Such a survey is a daunting prospect. Alternatively, we might choose to take sample of American females and estimate the average height in the population as the average height in the sample. How many should we include in the sample?

One possibility is to sample one person, chosen at random, and to use the height of that one person as an estimate of the average height in the population. The danger is self evident. There is reasonable chance that the person selected might be significantly smaller, or significantly taller than average. Thus there is a good chance we could be significantly off the mark. To guard against such mis-estimation, a larger sample may be chosen. If we chose 100 people at random, there may be some unusually small people, but there will also be some unusually tall people in the sample. These will cancel each other out when taking the average height in the sample and, hopefully, the average height in the sample will be fairly close to the true average height of the population. By increasing the sample size, the danger of mis-estimation has been reduced. The risk has been diversified. The risk of mis-estimation might be reduced even further by taking a larger sample.

In each of these examples, the idea of AVERAGING was important.

In the college grades example, the basic idea was that the AVERAGE grade on the tests, assignments and exams would reflect the expected ability of the student.

In the Stanley Cup the AVERAGING over seven games provided a "result per game played" that more closely reflected the abilities of the teams.

In the investment example, the usual measure of how well a portfolio performs is the rate of return. This is the dollar value of money earned in the portfolio divided by the amount of capital invested; i.e. the AVERAGE dollar return per dollar invested.

In the sampling example, the priniple was that the AVERAGE height in the sample provider a closer estimate of the true population height, the larger the size of the sample.

In each case, the basic idea of diversification was that as the numer of EVENTS (tests, hockey games, securities in the portfolio, or people in the sample) increased, so the average result provided a better reflection of the underlying (or expected) value

## A SIMPLE ILLUSTRATION OF DIVERSIFICATION

The process of diversification illustrated in the previous examples can be examined by setting up simple gambles. Perhaps the simplest of all gambles rests on the toss of a coin. There are only two possible outcomes, heads and tails. In addition, the odds are known in advance; there is a fifty percent chance of heads and a fifty percent chance of tails. To make the gamble interesting, we will place money on the outcome; heads, you win \$10 and tails you lose \$10. There is no bias in this gamble, you are just as likely to win as lose. The expected value is;

$$\text{Expected Value} = (0.5)(\$10) + (0.5)(-\$10) = \$0$$

If the coin is flipped just once, there are two possible outcomes. You may win \$10 or you may lose \$10. Even though the expected value is zero, there is no way you can achieve this outcome with just one toss of the coin. But now suppose that you toss the coin twice, with the \$10 stake resting on each toss. The possible outcomes are as follows;

TABLE 1

| First Toss | Second Toss | You Win | Average Win per toss | Probability |
|---|---|---|---|---|
| H | H | \$20 | \$10 | 1/4 |
| H | T | \$ 0 | \$ 0 | 1/4 |
| T | H | \$ 0 | \$ 0 | 1/4 |
| T | T | -\$20 | -\$10 | 1/4 |

Note; H = heads, T = tails

With the two tosses it is possible to break even. Either of the sequences "heads followed by tails" or "tails followed by heads" will result in a zero profit. But in addition, it is possible that you may win \$20 in total or lose \$20. There is a wider range of outcomes than with the single toss. It is useful to consider these outcomes on an averaged basis. This is achieved by dividing the total win or loss by the number of tosses. This column shows that the average win per toss may vary between \$10 and -\$10 but there is a fifty percent chance that it will be zero.

As with the illustrations given in the previous section, our concern here is with what happens to this AVERAGE outcome as the number of tosses (or events) increases. It is expected that the average outcome will become closer to the expected value as the number of tosses increases.

A simple experiment will reveal whether this does occur or not. A coin was tossed once and the result is shown in the first column of Table 2. The

result was Heads. On the same rules as the previous gamble, you win $10 with a Head. Thus averaging, you win gives $10 as shown in the final column. Then the coin was tossed twice. As luck would have it, there were two heads. The total win was therefore $20 which averages at $10 per toss as shown in the last column of the second row. The coin was tossed three times and the results are recorded in the third row. The coin was then tosses four times, five times and so on up to a sequence of twenty tosses. Now, for each sequence of tosses the EXPECTED VALUE of the average win per toss should be zero. We should find that the average value from the experiment gets closer to the expected value of zero as the number of tosses in the sequence increases.

TABLE 2

| Actual Sequence | # of Tosses | # of Heads | # of Tails | Average Win per toss $ |
|---|---|---|---|---|
| H | 1 | 1 | 0 | 10 |
| HH | 2 | 2 | 0 | 10 |
| TTT | 3 | 0 | 3 | -10 |
| HTHH | 4 | 3 | 1 | 5 |
| THTHT | 5 | 2 | 3 | - 2 |
| THHTTH | 6 | 3 | 3 | 0 |
| THHTHTH | 7 | 4 | 3 | 1.4 |
| THHTHHHH | 8 | 6 | 2 | 5 |
| TTHTTTTHT | 9 | 2 | 7 | - 5.6 |
| TTTTHHTTTH | 10 | 3 | 7 | - 4 |
| HTHTTHHHTTH | 11 | 6 | 5 | 0.9 |
| HTHTTTTHTTHH | 12 | 5 | 7 | - 1.7 |
| HHHHTHHTHTTTT | 13 | 7 | 6 | 0.8 |
| TTHHTTHHHHTTTT | 14 | 6 | 8 | - 1.4 |
| HHTHHHHTHHTTTTT | 15 | 8 | 7 | 0.7 |
| TTTTHTHHTTHTHTTH | 16 | 6 | 10 | - 2.5 |
| HHTHHTTTTTTHTHTHH | 17 | 8 | 9 | - 0.6 |
| THTHHHTHHTTTHHTHTT | 18 | 9 | 9 | 0 |
| THTHTTHTHTTHHHTTHTT | 19 | 8 | 11 | - 1.6 |
| HTHHTTTHTHHHTHHHTTTH | 20 | 11 | 9 | 1 |
| Total | 210 | 100 | 110 | - 0.476 |

The results of Table 2 do not fall into a perfect pattern. But neverthess, there does seem to be some pattern.

With a small sequence of tosses, e.g. between 1 and 5, there seems to be a good chance that the average win per toss will be far from the expected value of zero. The actual results were (10, 10. -10, 5 and 2).

With a larger sequence, e.g between 6 and 10 tosses, there appears to be a smaller chance that the average will differ so widely from zero (0, 1.4, 5, -5.6 and -4).

With a larger sequence still, e.g. between 11 and 15 tosses, the numbers appear to be grouped even close to zero (0.9, -1.7, 0.8, -1.4 and 0.7).

With an even larger sequence, e.g. the numbers, should be even more closely clustered about zero, (2.5, 0.6, 0, -1.6 and 1). In fact, we were thrown out a little by the effects of chance. The sequence of 16 was a little unkind by yielding too many heads.

Experiments such as these are somewhat vunerable to chance. The problem was that we only repeated each sequence just once. This is particularly troublesome for the small sequences. For example, consider that a sequence of two tosses were thrown followed by a sequence of four tosses. It is quite possible that the two toss sequence actually produced a head and a tail resulting in an average win of zero. In contrast, the four toss sequence could have turned up four heads resulting in an average win of $10. Such results would not help confirm the pattern of diversification. To get by the capricious effects of chance, so called RANDOM NOISE, we can make the experiment a little more sophisticated.

Now consider that the experiment shown in Table 2 is repeated, say 25 times. Thus there are;

25 sequences of one toss
25 sequences of two tosses
25 sequences of three tosses

etc

25 sequences of 20 tosses

The average win per toss, for say 'two toss sequences' will vary. Sometimes, the average may be $10, sometimes -$10, and sometimes zero. And similarly for each other sequence size, there will be 25 such averages. Table 3 plots out the whole set of results, comprising some 25 times 20 = 500 such averages. Thus column 1 of Table 3 is comparable to the final column of Table 2. It shows the average win per toss for games with 1, 2 3.... to 25 tosses per game. Of course the results are different because the coin landed differently. Each of the next 24 columns shows a similar set of results. The average win per toss is very variable for games in which the coin is tossed just one. This variation is seen in the first row and is measured by the standard deviation of 9.6. But when the games comprise a large number of tosses, e.g. the winnings are averaged over 20 tosses, the average win per toss is predictably close to zero. This is shown in the bottom row where the result of each game of 25 tosses is quite close to zero. Th standard deviation of these results is as 1.49.

Figure 1 summarizes the results of Table 3. The standard deviations and means of the average win per toss are shown. The values graphed are those in the last two columns of Table 3. The means of the various games are all quite close to zero reflected that the expected win from any game is zero. The standard deviation of the respective games fall dramatically as the number of tosses in the game is increased. In fact, the standard deviation is converging on zero. If we played a game in which a coin was tossed, say one

million times, the expected win per toss would still be zero, but the standard deviation of the win per toss would be but a hair's breadth from zero also.

An insurance pool is not too different from a casino. The casino repeats many bets with its customers. As shown in the previous section, its overall outcome is quite predictable when averaged over the large number of bets. The standard deviation of the average win per bet becomes smaller as the number of bets grows larger. In one sense, the insurance company makes a large number of "bets" with its policyholders. The same process of diversification implies that the insurance firm also will find that its average loss payment per policy becomes more predictable as the number of policies is increased. It will be seen that this is indeed the case if the right conditions are met. In the gambling analogy, successive tosses of a coin were unrelated or independent events. If insurance losses are independent events, then diversification works very well. But if insured losses are related or, more technically there is CORRELATION or COVARIANCE, the ability of the insurance company to diversify its risk is hampered.

As we proceed with this section we will be making further use of statistical concepts. These concepts, the expected value, variance, standard deviation and covariance have been used earlier. If you are still a little unsure about these ideas, it will be of advantage to review the statistical appendix at the end of the book before continuing with this chapter. These concepts will now be used to look at the expected value and risk of a portfolio.

## MEASURING THE EXPECTED VALUE AND RISK OF A PORTFOLIO

### (a). Defining a Portfolio: A Sum of Random Variable

A portfolio is a combination of risky outcomes. The Casino has a portfolio of bets and its owners are interested in the overall money outcome from its operations. Thus they will add or average the wins and losses to derive an overall measure of the performance of their operation. The averaging focusses attention away from individual bets. Thus the owners should not be unduly concerned if the occasional customer records a big win. In fact the prospect of a big win is probably necessary to attract a clientele. If the casino is well diversified, the losses to some customers will be more than covered by the gains from less fortunate "punters".

An investor may well hold a portfolio of financial securities such as stocks or bonds. The purpose, as we have seen earlier, is to average out any unusually bad performance on some securities with unusually good performance on others. The investors main focus of attention is on how well the portfolio as a whole performs. What rate of return is achieved per dollar of capital invested.

In both of these examples, a proper measure of success requires that individual outcomes be combined to give a measure of the overall performance of the portfolio. First, we will consider the expected value of a portfolio.

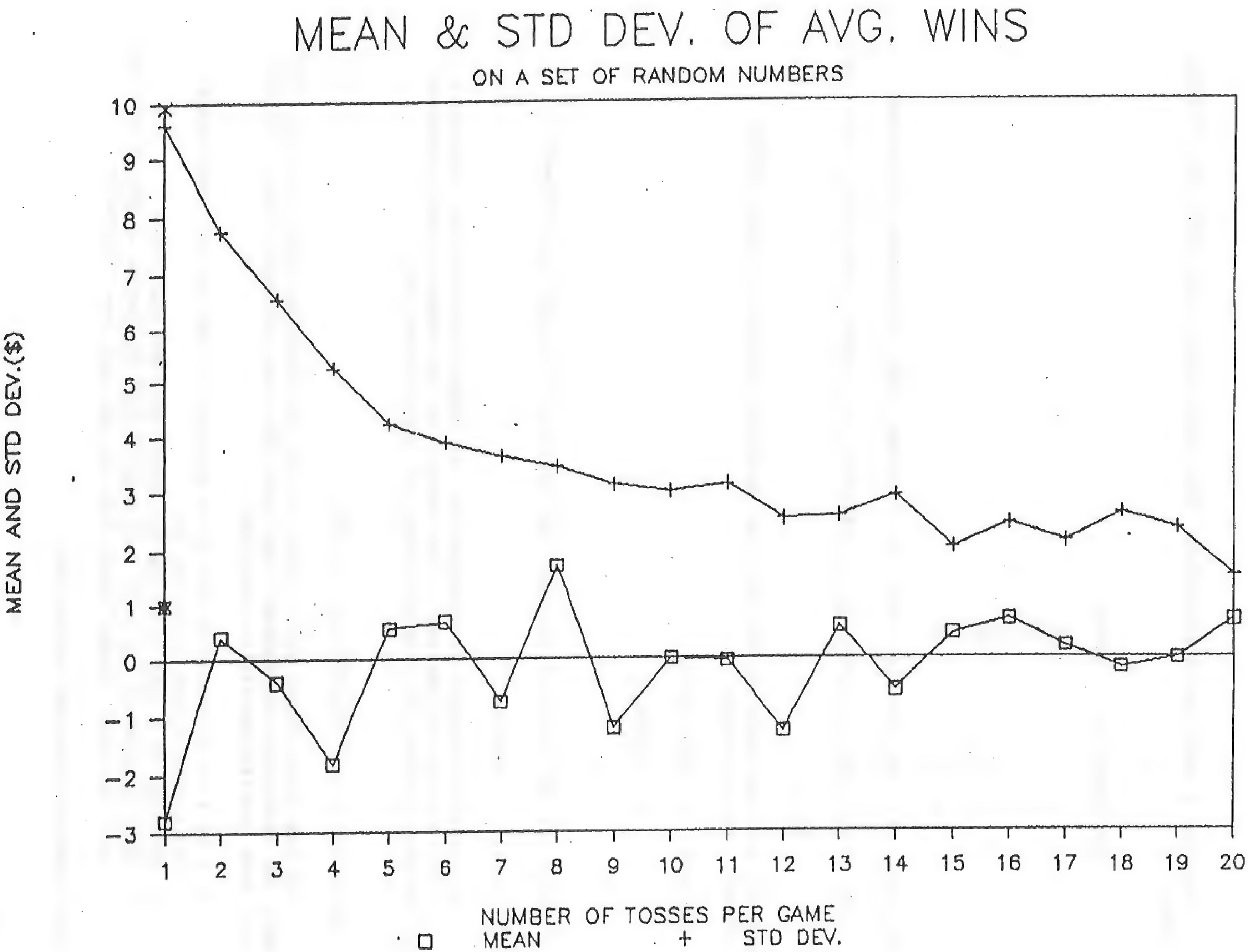
### (b). The Expected Value of a Portfolio

TABLE 3

MEANS AND STANDARD DEVIATIONS OF AVERAGE WIN($) PER TOSS

| GAMES | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | MEAN | STD DEV |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | (AVERAGE | VALUES | | OF | | WINS | PER | | TOSS) | | | | | | | | | |
| 1 | 10.0 | -10.0 | -10.0 | -10.0 | -10.0 | -10.0 | -10.0 | -10.0 | -10.0 | 10.0 | 10.0 | 10.0 | -10.0 | 10.0 | -10.0 | 10.0 | -10.0 | -10.0 | 10.0 | -10.0 | 10.0 | -10.0 | 10.0 | -10.0 | -10.0 | -2.80 | 9.60 |
| NUMBER 2 | 10.0 | -10.0 | -10.0 | 10.0 | 0.0 | 10.0 | 0.0 | 10.0 | 0.0 | 10.0 | 10.0 | 0.0 | -10.0 | 0.0 | -10.0 | -10.0 | 10.0 | 10.0 | -10.0 | -10.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.40 | 7.74 |
| OF 3 | -10.0 | -3.3 | 3.3 | 3.3 | -3.3 | 10.0 | 3.3 | -3.3 | 3.3 | 3.3 | -3.3 | -10.0 | -3.3 | -3.3 | -3.3 | 3.3 | 10.0 | -10.0 | 10.0 | -10.0 | -3.3 | -10.0 | 3.3 | 10.0 | 3.3 | -0.40 | 6.55 |
| TOSSES 4 | -5.0 | 0.0 | 0.0 | 0.0 | -5.0 | -10.0 | 5.0 | 5.0 | 0.0 | -5.0 | -10.0 | -10.0 | -5.0 | 5.0 | -5.0 | -10.0 | 0.0 | 0.0 | -5.0 | 0.0 | -5.0 | 0.0 | 10.0 | 5.0 | 0.0 | -1.80 | 5.27 |
| 5 | 2.0 | 6.0 | -2.0 | -2.0 | 2.0 | -6.0 | -2.0 | -6.0 | -2.0 | -2.0 | -2.0 | 2.0 | 6.0 | 6.0 | 2.0 | 6.0 | 6.0 | -2.0 | 6.0 | 6.0 | -6.0 | 2.0 | 2.0 | -6.0 | -2.0 | 0.56 | 4.22 |
| 6 | -10.0 | 0.0 | -3.3 | 3.3 | 6.7 | 3.3 | -3.3 | 6.7 | 0.0 | 3.3 | -3.3 | 0.0 | 6.7 | -3.3 | 3.3 | 0.0 | 3.3 | 3.3 | 3.3 | 0.0 | 0.0 | -3.3 | 3.3 | 0.0 | -3.3 | 0.67 | 3.69 |
| PER 7 | -7.1 | 4.3 | -1.4 | -4.3 | -7.1 | 1.4 | 1.4 | -7.1 | 1.4 | 4.3 | -4.3 | -1.4 | 4.3 | -1.4 | 1.4 | 1.4 | -4.3 | 4.3 | 1.4 | 4.3 | -1.4 | -1.4 | -1.4 | -4.3 | -1.4 | -0.74 | 3.64 |
| GAME 8 | -2.5 | 7.5 | 2.5 | 0.0 | 2.5 | 0.0 | -2.5 | 0.0 | 0.0 | 0.0 | 0.0 | -7.5 | 2.5 | 0.0 | 5.0 | 5.0 | 0.0 | 7.5 | 2.5 | 5.0 | 0.0 | 7.5 | 5.0 | 2.5 | 0.0 | 1.70 | 3.44 |
| 9 | -1.1 | 1.1 | -3.3 | -1.1 | -1.1 | 1.1 | -1.1 | 3.3 | -1.1 | -3.3 | 1.1 | -3.3 | -3.3 | 1.1 | -1.1 | -1.1 | -7.8 | -1.1 | -5.6 | -3.3 | -5.6 | 7.8 | 1.1 | 1.1 | -3.3 | -1.20 | 3.11 |
| 10 | 6.0 | -4.0 | -4.0 | -2.0 | 0.0 | 2.0 | 4.0 | 0.0 | -4.0 | 0.0 | -2.0 | -2.0 | 4.0 | -4.0 | 4.0 | 0.0 | -2.0 | 0.0 | 2.0 | -2.0 | -2.0 | -2.0 | 6.0 | 2.0 | 0.0 | 0.00 | 2.99 |
| 11 | 0.9 | -0.9 | -0.9 | -0.9 | -0.9 | -0.9 | 2.7 | -4.5 | -0.9 | -4.5 | -4.5 | 2.7 | 2.7 | -2.7 | 0.9 | 0.9 | 0.9 | -0.9 | 8.2 | 2.7 | -4.5 | 2.7 | 4.5 | -4.5 | 0.9 | -0.04 | 3.13 |
| NUMBER 12 | 0.0 | 0.0 | 0.0 | 1.7 | 0.0 | 0.0 | -1.7 | 3.3 | -8.3 | 0.0 | -1.7 | 0.0 | -5.0 | -3.3 | 0.0 | -3.3 | -1.7 | -1.7 | -1.7 | -1.7 | -1.7 | -5.0 | 0.0 | -3.3 | 3.3 | -1.27 | 2.51 |
| OF 13 | 3.8 | -0.8 | 5.4 | 0.8 | -2.3 | -0.8 | -0.8 | 5.4 | -0.8 | 5.4 | -2.3 | -2.3 | 3.8 | -2.3 | -3.8 | -0.8 | 0.8 | 0.8 | 0.8 | -0.8 | 0.8 | 2.3 | 2.3 | 0.8 | -0.8 | 0.58 | 2.55 |
| TOSSES 14 | -5.7 | -2.9 | 4.3 | 2.9 | -1.4 | -4.3 | 0.0 | -2.9 | -1.4 | 2.9 | -1.4 | -5.7 | -2.9 | 0.0 | -1.4 | 0.0 | -2.9 | 0.0 | -2.9 | 5.7 | -1.4 | 0.0 | -2.9 | -2.9 | 1.4 | -0.57 | 2.91 |
| 15 | -2.0 | 0.7 | 0.7 | 0.7 | 2.0 | -2.0 | 3.3 | 2.0 | 2.0 | 0.7 | 4.7 | -0.7 | -3.3 | -2.0 | 0.7 | -0.7 | 0.7 | 3.3 | -0.7 | -0.7 | -0.7 | 2.0 | -3.3 | 2.0 | 2.0 | 0.45 | 2.02 |
| 16 | 2.5 | 2.5 | 0.0 | 0.0 | 2.5 | 1.3 | 1.3 | 0.0 | 1.3 | 2.5 | 2.5 | -1.3 | 1.3 | 2.5 | -3.8 | 0.0 | -1.3 | 3.8 | 2.5 | 2.5 | -3.8 | -2.5 | -5.0 | 1.3 | 5.0 | 0.70 | 2.43 |
| PER 17 | -0.8 | 2.9 | -1.8 | -0.6 | 1.8 | 0.6 | -4.1 | -1.8 | 1.8 | 0.6 | -0.6 | 1.8 | -1.8 | 1.8 | 4.1 | -1.8 | -0.6 | 2.9 | -1.8 | -4.1 | -0.6 | 1.8 | 1.8 | 2.9 | 0.6 | 0.21 | 2.12 |
| GAME 18 | -1.1 | 1.1 | 5.6 | 2.2 | 2.2 | -1.1 | -2.2 | 0.0 | 3.3 | 4.4 | -1.1 | 2.2 | -1.1 | -2.2 | 2.2 | -2.2 | -3.3 | 0.0 | -3.3 | 0.0 | -4.4 | -4.4 | 1.1 | 0.0 | -2.2 | -0.18 | 2.60 |
| 19 | 0.5 | 0.5 | 3.7 | 1.6 | -0.5 | -0.5 | 3.7 | -3.7 | -3.7 | 1.6 | 2.6 | 1.6 | 0.5 | 0.5 | -2.6 | -0.5 | -0.5 | -2.6 | -1.6 | 0.5 | -5.8 | -0.5 | 0.5 | 3.7 | 0.5 | -0.02 | 2.33 |
| 20 | -1.0 | 2.0 | -1.0 | 2.0 | 0.0 | -1.0 | 3.0 | 0.0 | 3.0 | 2.0 | 0.0 | 0.0 | -1.0 | 1.0 | 1.0 | 2.0 | 1.0 | 3.0 | -3.0 | 1.0 | 0.0 | -1.0 | 1.0 | 2.0 | 1.0 | 0.68 | 1.49 |

FIGURE 1


MEAN & STD DEV. OF AVG. WINS
ON A SET OF RANDOM NUMBERS
MEAN AND STD DEV.($)
10
9
8
7
6
5
4
3
2
1
0
−1
−2
−3
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
NUMBER OF TOSSES PER GAME
MEAN
STD DEV.

Consider a very simple portfolio. You roll a die twice with the following prizes.

| Die Turns Up: | Prize |
|---|---|
| 1 | $1 |
| 2 | $2 |
| 3 | $3 |
| 4 | $4 |
| 5 | $5 |
| 6 | $6 |

For each roll, the expected value of the prize, E(X), is simply calculated;

$$E(X) = (1/6)(\$1) + (1/6)(\$2) + (1/6)(\$3) + (1/6)(\$4) + \$1/6)(\$5) + (1/6)(\$6)$$

$$\$3.5$$

But you have two rolls of the die. The portfolio outcome or total prize, E(Total), now is simply;

$$E(Total) = E(X) + E(X)$$

$$= (2)[E(X)]$$

$$= \$7$$

If there are "n" rolls of the die, the solution is simply generalised to;

$$E(Total) = (n)E(X) \tag{12.1}$$

The earlier examples concentrated on the AVERAGE outcome per occurrence, e.g. the average win per toss. The formula for the average win is obtained by dividing equation 12.1 by the number of rolls of the die, n;

$$E(Average\ Win) = [(n)E(X)]/n = E(X) \tag{12.2}$$

The same basic ideas may be used to find the expected value of a portfolio when the individual components do not have the same expected value. Supppose that your gambling portfolio comprised

1. A roll of the die with the above prizes .i.e the dollar prize equals the number turned up on the die
2. The toss of a coin with $10 for heads and -$10 for tails
3. The selection of one card from a pack with the prize of $26 if the card is a face card (King, Queen, Jack or Ace) and zero otherwise.

The respective expected values are;

$$E(P_1) \;=\; \$3.5$$

$$E(P_2) \;=\; \$0$$

$$E(P_3) \;=\; \$8 \qquad\qquad \text{(i.e. (16/52)(\$26) )}$$

The expected total value of the portfolio is;

$$\begin{aligned} E(X_1 + X_2 + X_3) &= E(X_1) + E(X_2) + E(X_3) \\ &= \$3.5 + \$0 + \$8 \\ &= \$11.5 \end{aligned}$$

A general formula for the total value is given together with a formula for the average win per event, each event being identified by the subscript "j";

$$E(\Sigma X) = \sum_j E(X_j) \tag{12.3}$$

$$= (n)E(X) \quad \text{if the events have identical expected values} \tag{12.4}$$

and

$$E(\Sigma X/n)$$

$$= \{ \sum_j E(X_j) \}/n \tag{12.5}$$

$$= E(X) \qquad \text{if the events have identical expected values} \tag{12.6}$$

(c). Measuring the Risk of a Portfolio

The general formula for measuring the standard deviation is;

$$\sigma = [\sum_i p_i (X_i - E(X))^2 ]^{1/2}. \tag{12.6}$$

With the rolling of the die, the expected standard deviation is:

$$\begin{aligned} \sigma^2 &= (1/6)(1 - 3.5)^2 + (1/6)(2 - 3.5)^2 + (1/6)(3 - 3.5)^2. \\ &+ (1/6)(4 - 3.5)^2 + (1/6)(5 - 3.5)^2 + (1/6)(6 - 3.5)^2. \\ &= \$4.18 \end{aligned}$$

But the present task is to calculate the standard deviation of the portfolio. Suppose, for example, the portfolio is simply the two rolls of the die. <u>The risk of the portfolio, will depend on the risk of each event. But it will also depend on the interaction, or COVARIANCE, between the events.</u> In

fact, the outcome of each roll of the die, is independent. If the first roll turns up, say six, there is no more, or no less likelihood that the second roll will turn up a six. Similarly, if you were to toss a coin five successive times and, by chance, gets five heads. The chances of getting a head on the sixth toss is still fifty percent. Coins and dies have no memory, they do not recall the results of previous events. Thus events such as rolling a die and tossing a coin are said to be INDEPENDENT; i.e. the COVARIANCE IS ZERO. This is not true of all events as we shall see later. But for the meantime, consider the risk in rolling a die twice. The standard deviation is the sum of the standard deviations from each event plus twice the covariance (which happens to be zero):

$$\sigma(X_1 + X_2) = [\ \sigma^2(X_1) + \sigma^2(X_2) + 2\ \mathrm{COV}(X_1;X_2)\ ]^{1/2}. \tag{12.8}$$

Which, in the present example gives the answer

$$\sigma(X_1 + X_2) = [\$4.18^2 + \$4.18^2 + 2(0)]^{1/2}.$$
$$= \$5.91$$

And the standard deviation of the average win is;

$$\sigma\{(X_1 + X_2)/2\} = \frac{[\sigma(X_1 + X_2]}{2} \tag{12.9}$$
$$= \frac{[\ \sigma^2(X_1) + \sigma^2(X_2) + 2\ \mathrm{COV}(X_1;X_2)\ ]^{1/2}.}{2}$$

Thus the rolling of the die twice produces a standard deviation averaged per event of;

$$\sigma\{(X_1 + X_2)/2\} = \frac{[4.18^2 + 4.18^2 + 2(0)]^{1/2}.}{2}$$
$$= 2.956$$

Equation 12.8 can be extended to portfolios that have, not two, but many individual events. The general formula is a little strange since it has a double summation sign for the covariances. We will explain this presently;

$$\sigma(\underset{i}{\Sigma X}) = [\ \underset{i}{\Sigma}\sigma_i^2 + \underset{i=j}{\Sigma\ \Sigma}\ \mathrm{COV}(Xi;Xj)\ ]^{1/2}. \tag{12.10}$$

And, the standard devation for the averaged result is

$$\sigma(\Sigma X/n) = \frac{\sigma(\Sigma X)}{n}$$

$$= \frac{[\, \sum_i \sigma_i^2 + \sum\sum_{i=j} COV(Xi;Xj)\,]^{1/2}}{n} \qquad (12.11)$$

What does equation 12.10 instruct us to do. It simply tells us to add up every variance and every combination of covariances. The first summation sign on the right side of equation 12.10 simply instructs that the variances be added. The second, double summation sign tells us that we should add all combinations of covariances between the events. The trick in using this formula is in knowing how many variances and covariances to look for. Perhaps the easiest way to see how many variances and covariances there may be is from the following Figure 12.2. Suppose that there are seven items in the portfolio. This means that there are seven variances. The V's in the diagonal of the matrix represent the seven variances. But now we have to consider every possible covariance. For example, we have the covariances between;

event 1 and event 2
event 1 and event 3
event 1 and event 4
......etc..........
event 2 and event 1
event 2 and event 3
event 2 and event 4
......etc..........
......etc..........
......etc..........
event 7 and event 4
event 7 and event 5
event 7 and event 6

Each of the squares in the matrix represents one such covariance. For example, the square marked with an "X" represents the covariance between item 4 and item 6.

FIGURE 12.2

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 1 | V | | | | | | |
| 2 | | V | | | | | |
| 3 | | | V | | | | |
| 4 | | | | V | | X | |
| 5 | | | | | V | | |
| 6 | | | | | | V | |
| 7 | | | | | | | V |

In all the seven by seven matrix has 49 cells; seven of these represent variances and the remaining 42 represent covariances. Now suppose that there had been 9 items in the portfolio. There would have been a 9 times 9 matrix with 81 elements. Of these, 9 would represent variances and the remaining 72 would represent covariances. In general, <u>with "n" items in the portfolio, there are n variances and n(n-1) covariances representing a total of n times n items to be included in equation 12.10</u>.

To illustrate the formulas, consider the earlier game of coin tossing. With a win of \$10 or -\$10 for a head or tail, the expected win and standard deviation for a single toss are;

$$E(X) = 0.5(\$10) + 0.5(-\$10)$$

$$= \$0$$

$$\sigma(X) = [\ 0.5(\$10 - \$0)^2 + 0.5(-\$10 - \$0)^2\ ]^{1/2}.$$

$$= \$10$$

Now, consider a game in which the coin is tossed 20 times. What is the expected win averaged per toss and the standard deviation of the win per toss? From equations 12.2 and 12.11

$$E(\Sigma X/n) = \$0 \qquad (12.2)$$

$$\sigma(\Sigma X/n) = \frac{[(20)10^2 + (20)(19)(0)]^{1/2}}{20}. \qquad (12.11)$$

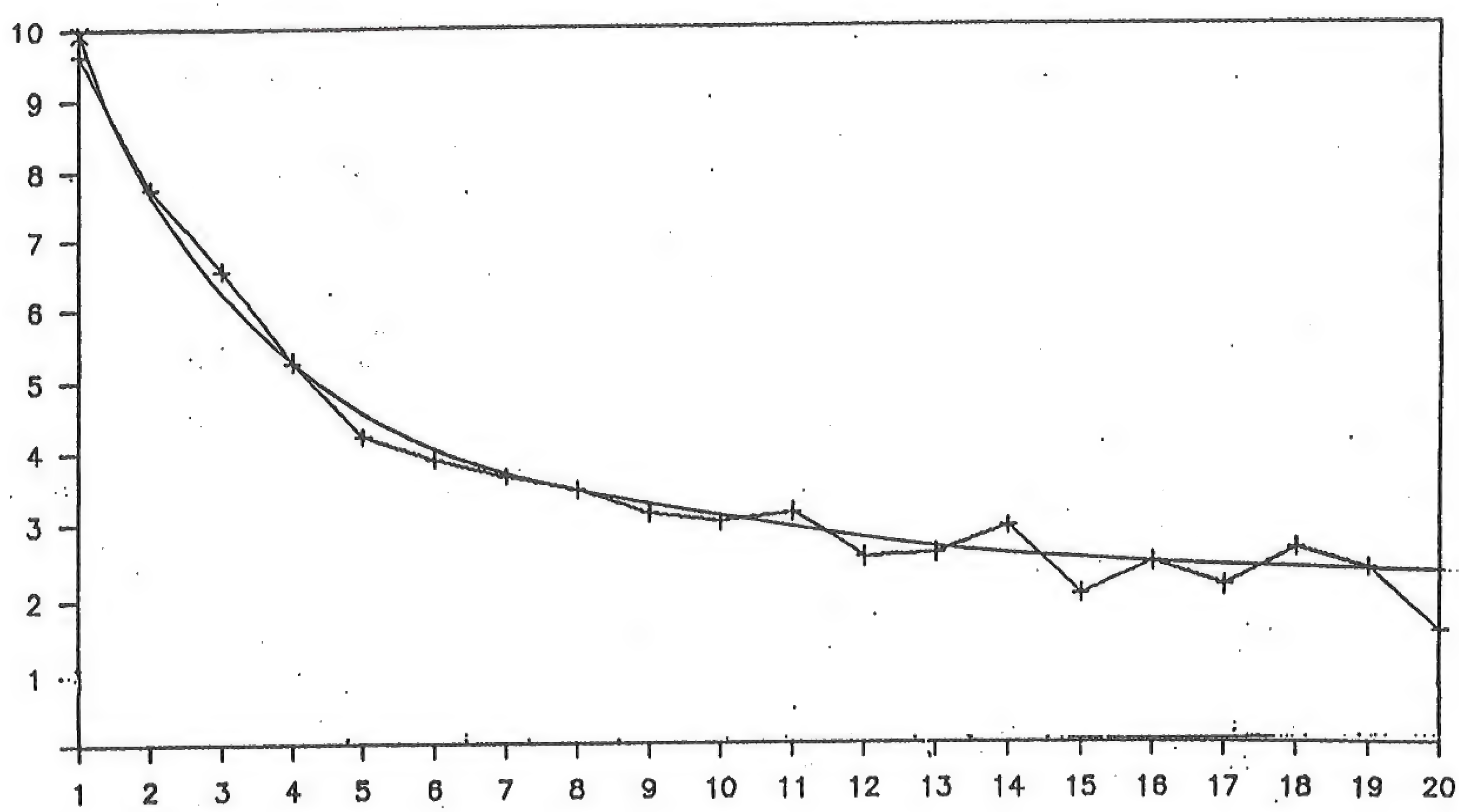
$$= 2.24$$

FIGURE 3

Standard Deviation of Average Win Per Toss

——— theoretical calculation from Equation 12.11

x----- calculated from random game presented in Figure 1



SCIENTIFIC AMERICAN



# Risk Analysis and Management

*Inadequate approaches to handling risks may result in bad policy. Fortunately, rational techniques for assessment now exist*

by M. Granger Morgan

Americans live longer and healthier lives today than at any time in their history. Yet they seem preoccupied with risks to health, safety and the environment. Many advocates, such as industry representatives promoting unpopular technology or Environmental Protection Agency staffers defending its regulatory agenda, argue that the public has a bad sense of perspective. Americans, they say, demand that enormous efforts be directed at small but scary-sounding risks while virtually ignoring larger, more commonplace ones.

Other evidence, however, suggests that citizens are eminently sensible about risks they face. Recent decades have witnessed precipitous drops in the rate and social acceptability of smoking, widespread shifts toward low-fat, high-fiber diets, dramatic improvements in automobile safety and the passage of mandatory seat belt laws—all steps that reduce the chance of untimely demise at little cost.

My experience and that of my colleagues indicate that the public can be very sensible about risk when companies, regulators and other institutions give it the opportunity. Laypeople have different, broader definitions of risk, which in important respects can be more rational than the narrow ones used by experts. Furthermore, risk management is, fundamentally, a question of values. In a democratic society, there is no acceptable way to make these choices without involving the citizens who will be affected by them.

The public agenda is already crowded with unresolved issues of certain or potential hazards such as AIDS, asbestos in schools and contaminants in food and drinking water. Meanwhile scientific and social developments are bringing new problems—global warming, genetic engineering and others—to the fore. To meet the challenge that these issues pose, risk analysts and managers will have to change their agenda for

M. GRANGER MORGAN has worked for many years to improve techniques for analyzing and managing risks to health, safety and the environment. Morgan heads the department of engineering and public policy at Carnegie Mellon University. He also holds appointments in the department of electrical and computer engineering and at the H. John Heinz III School of Public Policy and Management. Morgan received a B.A. from Harvard University, an M.S. from Cornell University and a Ph.D. in applied physics from the University of California, San Diego.

AIR DISASTER in Madrid claimed 183 lives in November 1983. The (small) chance of dying in an air crash is one of the prices that society agrees to pay for rapid, convenient global transportation. Some risks, including nuclear power generation, have caused fewer deaths but provoked greater calls for regulation, whereas others, such as automobiles, cause more deaths but arouse less concern.

evaluating dangers to the general welfare; they will also have to adopt new communication styles and learn from the populace rather than simply trying to force information on it.

While public trust in risk management has declined, ironically the discipline of risk analysis has matured. It is now possible to examine potential hazards in a rigorous, quantitative fashion and thus to give people and their representatives facts on which to base essential personal and political decisions.

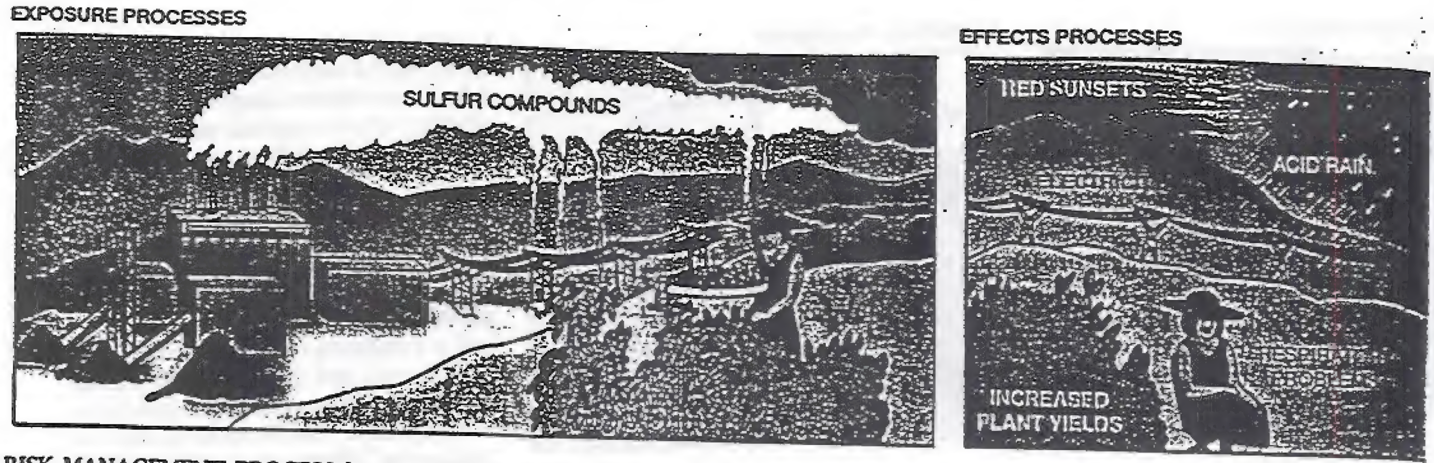
Risk analysts start by dividing hazards into two parts: exposure and effect. Exposure studies look at the ways in which a person (or, say, an ecosystem or a piece of art) might be subjected to change; effects studies examine what may happen once that exposure has manifested itself. Investigating the risks of lead for inner-city children, for example, might start with exposure studies to learn how old, flaking house paint releases lead into the environment and how children build up the substance in their bodies by inhaling dust or ingesting dirt. Effects studies might then attempt to determine the reduction in academic performance attributable to specific amounts of lead in the blood.

Exposure to a pollutant or other hazard may cause a complex chain of events leading to one of a number of effects, but analysts have found that the overall result can be modeled by a function that assigns a single number to any given exposure level. A simple, linear relation, for instance, accurately describes the average cancer risk incurred by smokers: 10 cigarettes a day generally increase the chance of contracting lung cancer by a factor of 25; 20 cigarettes a day increase it by a factor of 50. For other risks, however, a simple dose-response function is not appropriate, and more complex models must be used.

The study of exposure and effects is fraught with uncertainty. Indeed, uncertainty is at the heart of the definition of risk. In many cases, the risk may be well understood in a statistical sense but still be uncertain at the level of individual events. Insurance companies cannot predict whether any single driver will be killed or injured in an accident, even though they can estimate the annual number of crash-related deaths and injuries in the U.S. with considerable precision.

For other risks, such as those involving new technologies or those in which bad outcomes occur only rarely, uncertainty enters the calculations at a higher level—overall probabilities as well as individual events are unpredictable. If good actuarial data are not available, analysts must find other methods to estimate the likelihood of exposure and subsequent effects. The development of risk assessment during the past two decades has been in large part the story of finding ways to determine the extent of risks that have little precedent.

In one common technique, failure mode and effect analysis, workers try to identify all the events that might help cause a system to break down. Then they compile as complete a description




**RISK MANAGEMENT PROCESS** begins with analysis of the people and other entities exposed to change, such as in this illustration, from emissions from a coal-burning power plant (*left*). After the results of exposure have been quantified (*second panel*), they must then be filtered through public perceptions, which cause people to respond more strongly to some

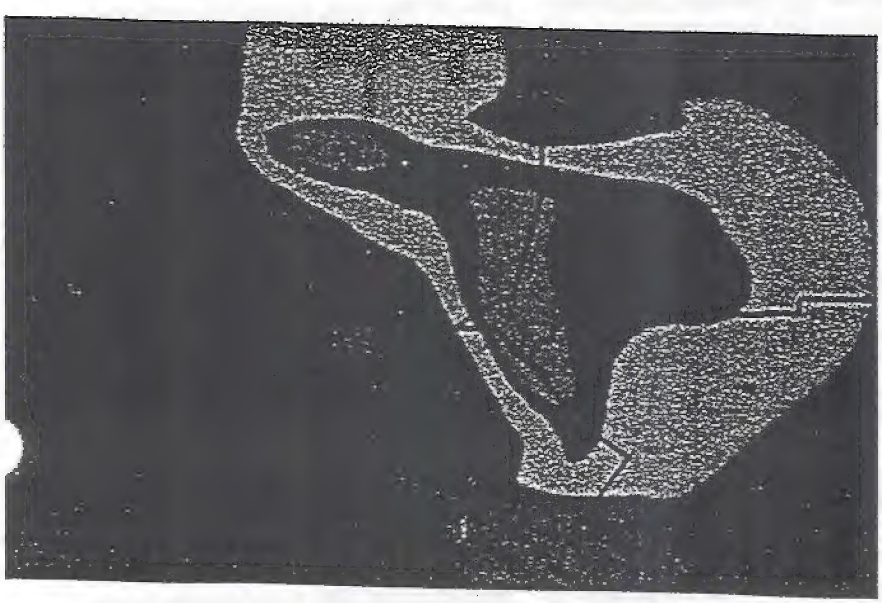
as possible of the routes by which those events could lead to a failure (for instance, a chemical tank might release its contents either because a weld cracks and the tank ruptures or because an electrical short causes the cooling system to stop, allowing the contents to overheat and eventually explode). Although enumerating all possible routes to failure may sound like a simple task, it is difficult to exhaust all the alternatives. Usually a system must be described several times in different ways before analysts are confident that they have grasped its intricacies, and even then it is often impossible to be sure that all avenues have been identified.

Once the failure modes have been enumerated, a fault tree can aid in estimating the likelihood of any given mode. This tree graphically depicts how the subsystems of an object depend on one another and how the failure of one part affects key operations. Once the fault tree has been constructed, one need only estimate the probability that individual elements will fail to find the chance that the entire system will cease to function under a particular set of circumstances. Norman C. Rasmussen of the Massachusetts Institute of Technology was among the first to use the method on a large scale when he directed a study of nuclear reactor safety in 1975. Although specific details of his estimates were disputed, fault trees are now used routinely in the nuclear industry and other fields.

Boeing applies fault-tree analysis to the design of large aircraft. Company engineers have identified and remedied a number of potential problems, such as vulnerabilities caused by routing multiple control lines through the same area. Alcoa workers recently used fault trees to examine the safety of their large furnaces. On the basis of their findings, the company revised its safety standards to mandate the use of programmable logic controllers for safety-critical controls. They also instituted rigorous testing of automatic shut-off valves for leaks and added alarms that warn operators to close manual isolation valves during shutdown periods. The company estimates that these changes have reduced the likelihood of explosions by a factor of 20. Major chemical companies such as Du Pont, Monsanto and Union Carbide have also employed the technique in designing processes for chemical plants, in deciding where to build plants and in evaluating the risks of transporting chemicals.

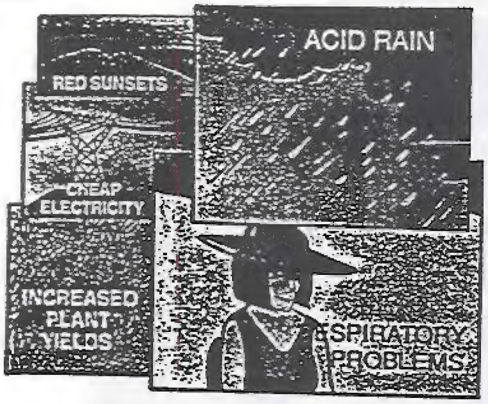
In addition to dealing with uncertainty about the likelihood of an event such as the breakdown of a crucial piece of equipment, risk analysts must cope with other unknowns: if a chemical tank leaks, one cannot determine beforehand the exact amount of pollutant released, the precise shape of the resulting dose-response curves for people exposed, or the values of the rate constants governing the chemical reactions that convert the contents of the tank to more or less dangerous forms. Such uncertainties are often represented by means of probability distributions, which describe the odds that a quantity will take on a specific value within a range of possible levels.

When risk specialists must estimate the likelihood that a part will fail or assign a range of uncertainty to an essential value in a model, they can sometimes use data collected from similar systems elsewhere—although the design of a proposed chemical plant as a whole may be new, the components in its high-pressure steam systems will ba-

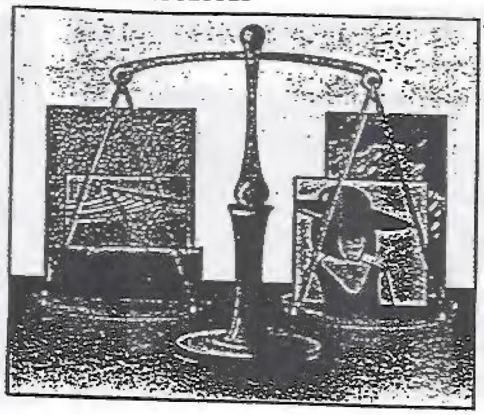


**SUPERCOMPUTER MODEL** of ozone concentrations in the Los Angeles basin (*pink, highest; yellow, lowest*) serves as a starting point for analyses of the risks of exposure to air pollutants.

PERCEPTION PROCESSES




VALUATION PROCESSES



aspects of risk than to others. Ultimately, costs and benefits will be weighed. Agreeing on the values used to make decisions and making sure that all relevant effects are taken into account are crucial, but often neglected, parts of the process.

sically be indistinguishable from those in other plants.

In other cases, however, historical data are not available. Sometimes workers can build predictive models to estimate probabilities based on what is known about roughly similar systems, but often they must rely on expert subjective judgment. Because of the way people think about uncertainty, this approach may involve serious biases. Even so, quantitative risk analysis retains the advantage that judgments can be incorporated in a way that makes assumptions and biases explicit.

Only a few years ago such detailed study of risks required months of custom programming and days or weeks of mainframe computer time. Today a variety of powerful, general-purpose tools are available to make calculations involving uncertainty. These programs, many of which run on personal computers, are revolutionizing the field. They enable accomplished analysts to complete projects that just a decade ago were considered beyond the reach of all but the most sophisticated organizations [*see box on page 38*]. Although using such software requires training, they could democratize risk assessment and make rigorous determinations far more widely available.

After they have determined the likelihood that a system could expose people to harm and described the particulars of the damage that could result from exposure, some risk analysts believe their job is almost done. In fact, they have just completed the preliminaries. Once a risk has been identified and analyzed, psychological and social processes of perception and valuation come into play. How people view and evaluate particular risks determines which of the many changes that may occur in the world they choose to notice and perhaps do something about. Someone must then establish the rules for weighing risks, for deciding if the risk is to be controlled and, if so, how. Risk management thus tends to force a society to consider what it cares about and who should bear the burden of living with or mitigating a problem once it has been identified.

For many years, most economists and technologists perceived risk simply in terms of expected value. Working for a few hours in a coal mine, eating peanut butter sandwiches every day for a month, and living next to a nuclear power plant for five years all involve an increased risk of death of about one in a million, so analysts viewed them all as equally risky. When people are asked to rank various activities and technologies in terms of risk, however, they produce lists whose order does not correspond very closely to the number of expected deaths. As a result, some early risk analysts decided that people were confused and that their opinions should be discounted.

Since then, social scientists have conducted extensive studies of public risk perception and discovered that the situation is considerably more subtle. When people are asked to order well-known hazards in terms of the number of deaths and injuries they cause every year, on average they can do it pretty well. If, however, they are asked to rank those hazards in terms of risk, they produce quite a different order.

People do not define risk solely as the expected number of deaths or injuries per unit time. Experimental psychologists Baruch Fischhoff of Carnegie Mellon University and Paul Slovic and Sarah Lichtenstein of Decision Research in Eugene, Ore., have shown that people also rank risks based on how well the process in question is understood, how equitably the danger is distributed, how well individuals can control their exposure and whether risk is assumed voluntarily.

Slovic and his colleagues have found that these factors can be combined into three major groups. The first is basically an event's degree of dreadfulness (as determined by such features as the scale of its effects and the degree to which it affects "innocent" bystanders). The second is a measure of how well the risk is understood, and the third is the number of people exposed. These groups of characteristics can be used to define a "risk space." Where a hazard falls within this space says quite a lot about how people are likely to respond to it. Risks carrying a high level of "dread," for example, provoke more calls for government intervention than do some more workaday risks that actually cause more deaths or injuries.

In making judgments about uncertainty, including ones about risk, experimental psychologists have found that people unconsciously use a number of heuristics. Usually these rules of thumb work well, but under some circumstances they can lead to systematic bias or other errors. As a result, people tend to underestimate the frequency of very common causes of death—stroke, cancer, accidents—by roughly a factor of 10. They also overestimate the frequency of very uncommon causes of death (botulism poisoning, for example) by as much as several orders of magnitude.

These mistakes apparently result from the so-called heuristic of availability. Daniel Kahneman of the University of California at Berkeley, Amos N. Tversky of Stanford University and others have found that people often judge the likelihood of an event in terms of how easily they can recall (or imagine) examples. In this case, stroke is a very common cause of death, but most people learn about it only when a close friend or relative or famous person dies; in contrast, virtually every time someone dies of botulism, people are likely to hear about it on the evening news. This heuristic and others are not limited to the general public. Even experts sometimes employ them in making judgments about uncertainty.

Once people have noticed a risk and decided that they care enough to do something about it, just what should they do? How should they decide the amount to be spent on reducing the risk, and on whom should they place the primary burdens? Risk managers can intervene at many points: they can work to prevent the process producing the risk, to reduce exposures,

to modify effects, to alter perceptions or valuations through education and public relations or to compensate for damage after the fact. Which strategy is best depends in large part on the attributes of the particular risk.

Even before determining how to intervene, risk managers must choose the rules that will be used to judge whether to deal with a particular issue and, if so, how much attention, effort and money to devote. Most rules fall into one of three broad classes: utility based, rights based and technology based. The first kind of rules attempt to maximize net benefits. Analysts add up the pros and cons of a particular course of action and take the difference between the two. The course with the best score wins.

Early benefit-cost analyses employed fixed estimates of the value of good and bad outcomes. Many workers now use probabilistic estimates instead to reflect the inherent uncertainty of their descriptions. Although decisions are ultimately made in terms of expected values, other measures may be employed as well. For example, if the principal concern is to avoid disasters, analysts could adopt a "minimax" criterion, which seeks to minimize the harm done by the worst possible outcome, even if that leads to worse results on average.

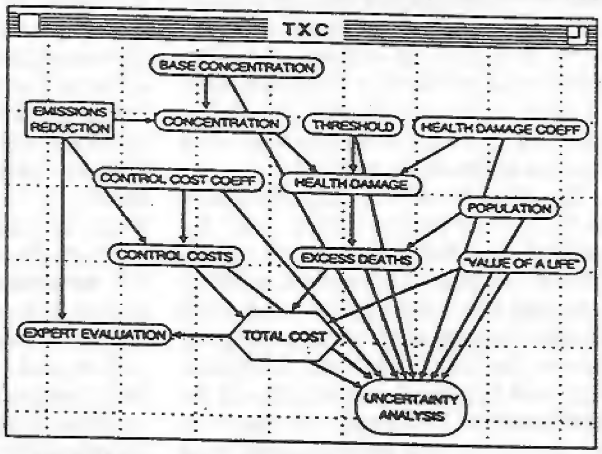
Of course, many tricky points are involved in such calculations. Costs and benefits may not depend linearly on the amount of pollutant emitted or on the number of dollars spent for control. Furthermore, not all the pros and cons of an issue can necessarily be measured on the same scale. When the absolute magnitude of net benefits cannot be estimated, however, rules based on relative criteria such as cost-effectiveness can still aid decision makers.

Rights-based rules replace the notion of utility with one of justice. In most utility-based systems, anything can be subject to trade-offs; in rights-based ones, however, there are certain things that one party cannot do to another without its consent, regardless of costs or benefits. This is the approach that Congress has taken (at least formally) in the Clean Air Act of 1970: the law does not call for maximizing net social benefit; instead it just requires controlling pollutant concentrations so as to protect the most sensitive populations exposed to them. The underlying pre-

## Risk Analysis in Action

Uncertainty is a central element of most problems involving risk. Analysts today have a number of software tools that incorporate the effects of uncertainty. These tools can show the logical consequences of a particular set of risk assumptions and rules for making decisions about it. One such system is Demos, developed by Max Henrion of Lumina Decision Systems in Palo Alto, Calif.

To see how the process works, consider a hypothetical chemical pollutant, "TXC." To simplify matters, assume that the entire population at risk (30 million people) is exposed to the same dose—this makes a model of exposure processes unnecessary. The next step is to construct a function that describes the risk associated with any given exposure level—for example, a linear dose-response function, possibly with a threshold below which there is no danger.




BLOCKS in the diagram above can be expanded to call up a window containing graphs and tables for their assumptions, equations and probability distributions.

Given this information, Demos can estimate the number of excess deaths caused every year by TXC exposure. According to the resulting cumulative probability distribution, there is about a 30 percent chance that no one dies, about a 50 percent chance that fewer than 100 people die each year and about a 10 percent chance that more than 1,000 die.

Meanwhile, for a price, pollution controls can reduce the concentration of TXC. (The cost of achieving any given reduction, like the danger of exposure, is determined by consultation with experts.) To choose a level of pollution control that minimizes total social costs, one must first decide how much society is willing to invest to prevent mortality. The upper and lower bounds in this example are $300,000 and $3 million per death averted. (Picking such numbers is a value judgment; in practice, a crucial part of the analysis would be to find out how sensitive the results are to the dollar values placed on life or health.)

Net social costs, in this model, are simply the sum of control costs and mortality. At $300,000 per death averted, their most likely value reaches a minimum when TXC emissions are reduced by 55 percent. At $3 million, the optimum reduction is about 88 percent.

Demos can also calculate a form of correlation between each of the input variables and total costs. Strong correlations indicate variables that contribute significantly to the uncertainty in the final cost estimate. At low levels of pollution control, possible variations in the slope of the damage function, in the location of the threshold and in the base concentration of the pollutant contribute the most to total uncertainty. At very high levels of control, in contrast, almost all the uncertainty derives from unknowns in the cost of controlling emissions.

Finally, Demos can compute the difference in expected cost between the optimal decision based on current information and that given perfect information—that is, the benefit of removing all uncertainties from the calculations. This is known in decision analysis as the expected value of perfect information; it is an upper bound on the value of research. If averting a single death is worth $300,000 to society, this value is $38 million a year; if averting a death is worth $3 million, it is $71 million a year.

Although tools such as Demos put quantitative risk analysis within reach of any group with a personal computer, using them properly requires substantial education. My colleagues and I found that a group of first-year engineering doctoral students first exposed to Demos tended to ignore possible correlations among variables, thus seriously overestimating the uncertainty of their results.

sumption holds that these individuals have a right to protection from harm.

Technology-based criteria, in contrast to the first two types, are not concerned with costs, benefits or rights but rather with the level of technology available to control certain risks. Regulations based on these criteria typically mandate "the best available technology" or emissions that are "as low as reasonably achievable." Such rules can be difficult to apply because people seldom agree on the definitions of "available" or "reasonably achievable." Furthermore, technological advances may impose an unintended moving target on both regulators and industry.

There is no correct choice among the various criteria for making decisions about risks. They depend on the ethical and value preferences of individuals and society at large. It is, however, critically important that decision frameworks be carefully and explicitly chosen and that these choices be kept logically consistent, especially in complex situations. To do otherwise may produce inconsistent approaches to the same risk. The EPA has slipped into this error by writing different rules to govern exposure to sources of radioactivity that pose essentially similar risks.

Implicit in the process of risk analysis and management is the crucial role of communication. If public bodies are to make good decisions about regulating potential hazards, citizens must be well informed. The alternative of entrusting policy to panels of experts working behind closed doors has proved a failure, both because the resulting policy may ignore important social considerations and because it may prove impossible to implement in the face of grass-roots resistance.

Until the mid-1980s, there was little research on communicating risks to the public. Over the past five years, along with my colleagues Fischhoff and Lester B. Lave, I have found that much of the conventional wisdom in this area does not hold up. The chemical industry, for example, distilled years of literature about communication into advice for plant managers on ways to make public comparisons between different kinds of risks. We subjected the advice to empirical evaluation and found that it is wrong. We have concluded that the only way to communicate risks reliably is to start by learning what people already know and what they need to know, then develop messages, test them and refine them until surveys demonstrate that the messages have conveyed the intended information.

In 1989 we looked at the effects of the EPA's general brochure about radon in homes. The EPA prepared this brochure according to traditional methods: ask scientific experts what they think people should be told and then package the result in an attractive form. In fact, people are rarely completely ignorant about a risk, and so they filter any message through their existing knowledge. A message that does not take this filtering process into account can be ignored or misinterpreted.

To study people's mental models, we began with a set of open-ended interviews, first asking, "Tell me about radon." Our questions grew more specific only in the later stages of the interview. The number of new ideas encountered in such interviews approached an asymptotic limit after a couple of dozen people. At this point, we devised a closed-form questionnaire from the results of the interviews and administered it to a much larger sample.

We uncovered critical misunderstandings in beliefs that could undermine the effectiveness of the EPA's messages. For example, a sizable proportion of the public believes that radon contamination is permanent and does not go away. This misconception presumably results from an inappropriate inference based on knowledge about chemical contaminants or long-lived radioisotopes. The first version of the EPA's "Citizen's Guide to Radon" did not discuss this issue. Based in part on our findings, the latest version addresses it explicitly.

The objective of risk communication is to provide people with a basis for making an informed decision; any effective message must contain information that helps them in that task. With former doctoral students Ann Bostrom, now at the Georgia Institute of Technology, and Cynthia J. Atman, now at the University of Pittsburgh, we used our method to develop two brochures about radon and compared their effectiveness with that of the EPA's first version. When we asked people to recall simple facts, they did equally well with all three brochures. But when faced with tasks that required inference—advising a neighbor with a high radon reading on what to do—people who received our literature dramatically outperformed those who received the EPA material.

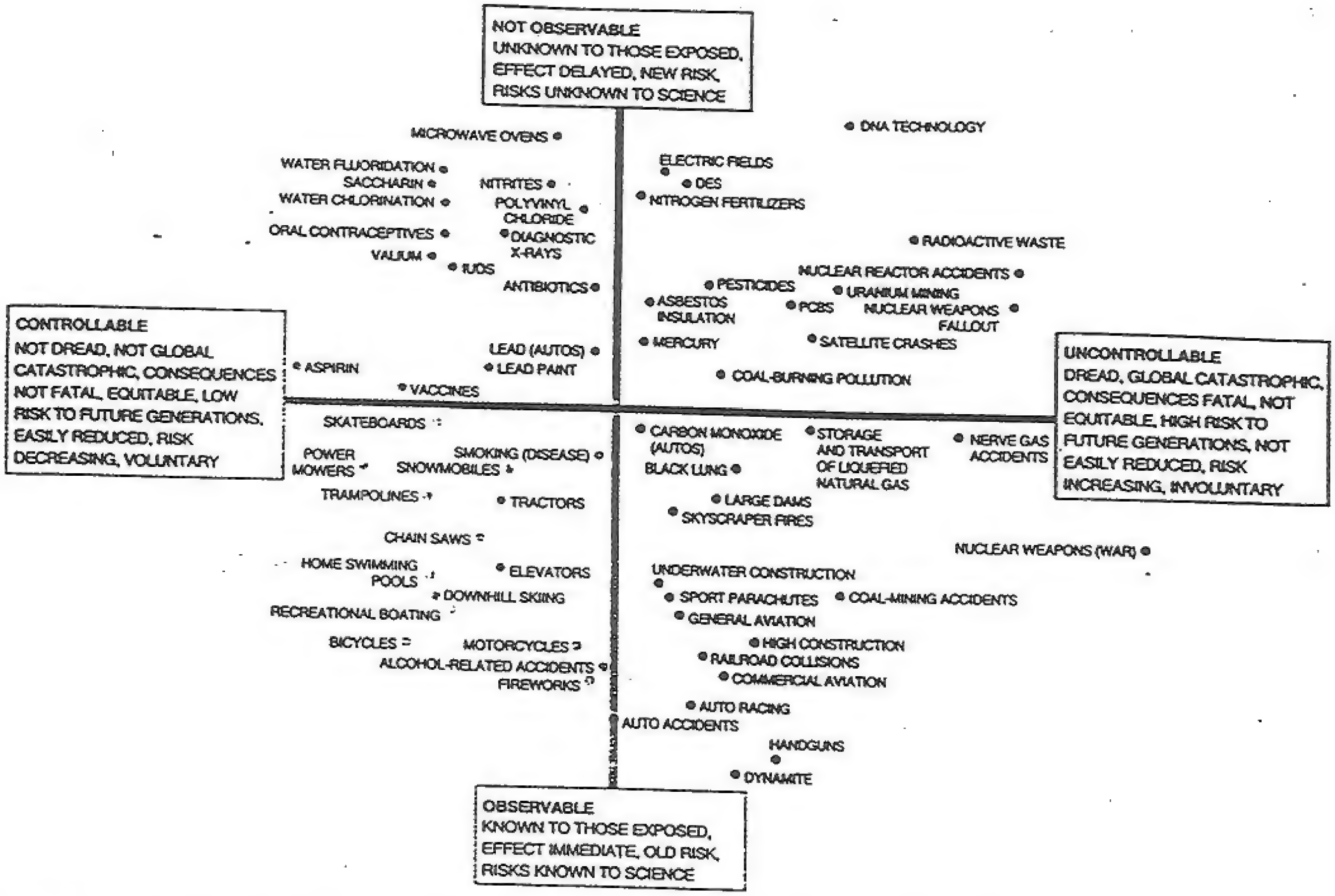
We have found similar misperceptions in other areas, say, climatic change. Only a relatively small proportion of people associate energy use and carbon dioxide emissions with global warming. Many believe the hole in the ozone layer is the factor most likely to lead to global warming, although in fact the two issues are only loosely connected. Some also think launches of spacecraft are the major contributor to holes in the ozone layer. (Willett Kempton of the University of Delaware has found very similar perceptions.)

The essence of good risk communication is very simple: learn what people already believe, tailor the communication to this knowledge and to the decisions people face and then subject the resulting message to careful empirical evaluation. Yet almost no one communicates risks to the public in this fashion. People get their information in fragmentary bits through a press that often does not understand technical details and often chooses to emphasize the sensational. Those trying to convey information are generally either advocates promoting a particular agenda or regulators who sometimes fail either to do their homework or to take a sufficiently broad perspective on the risks they manage. The surprise is not that opinion on hazards may undergo wide swings or may sometimes force silly or inefficient outcomes. It is that the public does as well as it does.

Indeed, when people are given balanced information and enough time to reflect on it, they can do a remarkably good job of deciding what problems are important and of systematically addressing decisions about risks. I conducted studies with Gordon Hester (then a doctoral student, now at the Electric Power Research Institute) in which we asked opinion leaders—a teacher, a state highway patrolman, a bank manager and so on—to play the role of a citizens' board advising the governor of Pennsylvania on the siting of high-voltage electric transmission lines. We asked the groups to focus particularly on the controversial problem of health risks from electric and magnetic fields emanating from transmission lines. We gave them detailed background information and a list of specific questions. Working mostly on their own, over a period of about a day and a half (with pay), the groups structured policy problems and prepared advice in a fashion that would be a credit to many consulting firms.

If anyone should be faulted for the poor quality of responses to risk, it is probably not the public but rather risk managers in government and industry. First, regulators have generally adopted a short-term perspective focused on taking action quickly rather than investing in the research needed to improve understanding of particular hazards in the future. This focus is especially evident in regulations that have been formulated to ensure the safety of the environment, workplace and consumer products.

Second, these officials have often




RISK SPACE has axes that correspond roughly to a hazard's "dreadfulness" and to the degree to which it is understood. Risks in the upper right quadrant of this space are most likely to provoke calls for government regulation.

adopted too narrow an outlook on the risks they manage. Sometimes attempts to reduce one risk (burns from flammable children's pajamas) have created others (the increased chance of cancer from fireproofing chemicals).

In some instances, regulators have ignored large risks while attacking smaller ones with vigor. Biologist Bruce Ames of Berkeley has argued persuasively that government risk managers have invested enormous resources in controlling selected artificial carcinogens while ignoring natural ones that may contribute far more to the total risk for human cancer.

Third, government risk managers do not generally set up institutions for learning from experience. Too often adversarial procedures mix attempts to figure out what has happened in an incident with the assignment of blame. As a result, valuable safety-related insights may either be missed or sealed away from the public eye. Civilian aviation, in contrast, has benefited extensively from accident investigations by the National Transportation Safety Board. The board does its work in isolation from arguments about liability; its results are widely published and have contributed measurably to improving air safety.

Many regulators are probably also too quick to look for single global solutions to risk problems. Experimenting with multiple solutions to see which ones work best is a strategy that deserves far more attention than it has received. With 50 states in a federal system, the U.S. has a natural opportunity to run such experiments.

Finally, risk managers have not been sufficiently inventive in developing arrangements that permit citizens to become involved in decision making in a significant and constructive way, working with experts and with adequate time and access to information. Although there are provisions for public hearings in the licensing process for nuclear reactors or the siting of hazardous waste repositories, the process rarely allows for reasoned discussion, and input usually comes too late to have any effect on the set of alternatives under consideration.

Thomas Jefferson was right: the best strategy for assuring the general welfare in a democracy is a well-informed electorate. If the U.S. and other nations want better, more reasoned social decisions about risk, they need to take steps to enhance public understanding. They must also provide institutions whereby citizens and their representatives can devote attention to risk management decisions. This will not preclude the occasional absurd outcome, but neither does any other way of making decisions. Moreover, appropriate public involvement should go a long way toward eliminating the confrontational tone that has become so common in the risk management process.


FURTHER READING

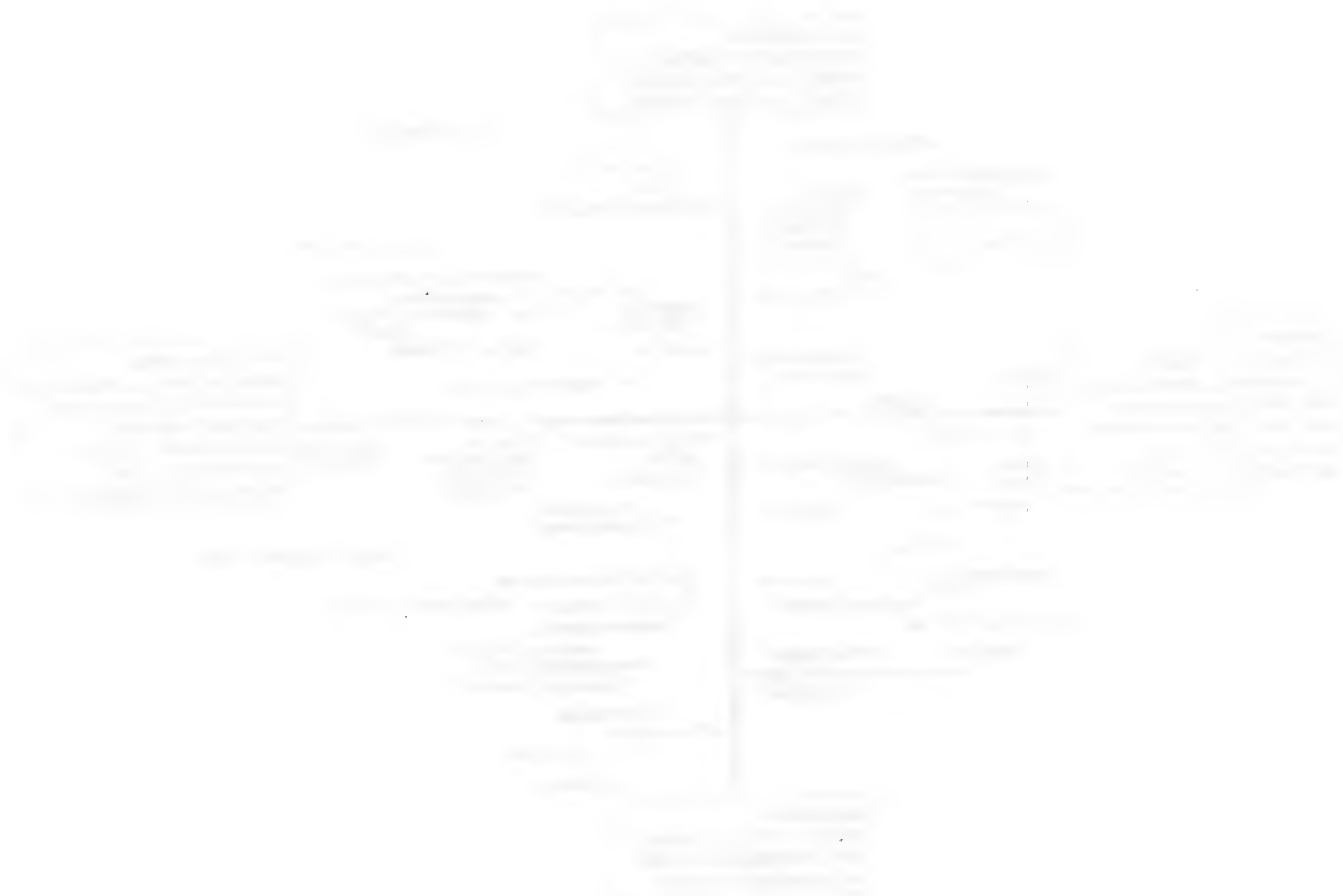
RATIONAL CHOICE IN AN UNCERTAIN WORLD. Robyn M. Dawes. Harcourt Brace Jovanovich, 1988.

READINGS IN RISK. Edited by Theodore S. Glickman and Michael Gough. Resources for the Future, 1990.

UNCERTAINTY: A GUIDE TO DEALING WITH UNCERTAINTY IN QUANTITATIVE RISK AND POLICY ANALYSIS. M. Granger Morgan and Max Henrion. Cambridge University Press, 1990.

COMMUNICATING RISK TO THE PUBLIC. M. Granger Morgan, Baruch Fischhoff, Ann Bostrom, Lester Lave and Cynthia J. Atman in *Environmental Science and Technology*, Vol. 26, No. 11, pages 2048–2056; November 1992.

RISK ANALYSIS. Publication of the Society for Risk Analysis, published quarterly by Plenum Publishing.



Midland Corporate Finance Journal
#3, No 2 Summer 1985



6

# An Integrated Approach to Corporate Risk Management

Alan C. Shapiro and Sheridan Titman,
University of California at Los Angeles

See also Greenwald & Stiglitz pp16 -- AER May 1990



In determining their risk management policy, many firms solicit the opinions of investment bankers and commercial lenders. They also pay attention to the yields and ratings on their debt since this gives them direct evidence on how the market perceives their current risk profile. Firms also rely on industry standards in deciding how much debt to issue and how much insurance to purchase.

These approaches are useful because they give the firm access to potentially valuable information about what a "normal" risk profile would look like. But relying too heavily on the opinions of lenders or industry standards doesn't take into account the unique attributes of a particular firm. In many cases, there are no firms whose business breakdown and product market strategy are similar enough to be directly comparable. Even where there is a well-defined industry, this doesn't mean that the risk profiles of other companies are optimal for their current circumstances. For example, there is some evidence that individual debt ratios are, to a large extent, a product of random historical events.[1] Moreover, the potential costs of financial distress to a firm may be quite large, even if the risk to a lender is not that great. If these costs of financial distress are high enough, the fact that, for example, lenders are willing to provide additional funds doesn't mean that the firm should borrow more.

Corporate managements seem to recognize the costs associated with a high risk profile because they engage in a wide variety of risk-reducing behavior. They buy insurance to protect against property and casualty losses and product liability suits. They use commodity and financial futures and forward contracts to guard against fluctuations in interest rates, foreign exchange rates, and the prices of specific products. They shun risky products, sometimes even when the promised returns are high. And, in many cases, they restrict the amount of financial leverage they employ, even when there may be substantial tax advantages to borrowing.

Typically, these decisions—such as how much fire insurance to buy, whether to hedge a particular foreign exchange risk, and how much leverage to incorporate within the company's capital structure—are made independently of one another, presumably because each deals with a different source of risk. But because each of these decisions affects the total risk of the firm (albeit with different costs and consequences), there are clearly benefits to integrating risk management activities into a single framework. To do this properly, however, requires answers to two major questions: (1) What factors should management consider in deciding the firm's optimal risk profile? (2) What are the relevant tradeoffs involved in choosing among the various risk-reducing or hedging mechanisms available? (For example, should one reduce corporate risk by lowering the debt-equity ratio or by taking out a larger product liability policy?)

Unfortunately, modern financial theory offers little guidance in such matters. Indeed, the theory of risk in modern finance, as embodied by the capital asset pricing model (CAPM) and the more recent

1. Some academics have suggested that corporate capital structures are determined in large part by companies' current investment requirements and their past profitability (and thus the availability of internal funds); to the extent this is true, actual debt ratios may be more the product of random economic events than any conscious management adherence to a capital structure "target." See Stewart C. Myers, "The Capital Structure Puzzle," *Journal of Finance* (July, 1984), pp.575-592; also Sheridan Titman and Roberto Wessels, "The Determinants of the Capital Structure Choice," UCLA working paper, 1984.

arbitrage pricing theory (APT), seems to regard as irrelevant, if not actually wasteful, a range of corporate hedging activities designed to reduce the total risk, or variability, of the firm's cash flows. Both the CAPM and the APT demonstrate that, under reasonable circumstances, diversifiable risks are not "priced" by sophisticated investors and, hence, do not affect the stock market's required rates of return. Systematic or "market" risks (those which cannot be diversified away by investors) are priced; but because the price of risk is the same for all market participants, there is no gain to shareholders from "laying them off" to financial markets. Consequently, as this reasoning goes, the expected net present value of buying insurance or a futures or forward contract should be zero in an efficient market. In this light, management decisions to insure or hedge assets appear, at best, "neutral mutations" (having no effect on the value of the firm). At worst, such actions, to the extent they are costly, are viewed as "irrational behavior" penalizing corporate stockholders.

The purpose of this article is to present and expand upon a relatively new justification for corporate hedging practices—"new," that is, in the academic finance literature. We begin by offering a rationale for actively managing *total* corporate risk that is consistent with both the premise of shareholder wealth maximization and the exclusive focus of asset pricing models on systematic risk. We then go on to use this theoretical framework to generate a set of principles to guide corporate management in establishing a coherent, centralized approach to risk management. Such an approach considers the costs and benefits of a variety of available risk-reducing tools and strategies for managing the total exposure of the company.

## Why Total Risk Matters

Modern finance theory holds that the value of a firm is equal to its expected future cash flows discounted at the appropriate interest rate. Financial economists have concerned themselves almost exclusively with the effect of risk on market discount rates, for the most part ignoring its effect on expected cash flow. According to both the CAPM and APT, sophisticated investors require higher rates of return on securities imposing greater risk; but because such investors diversify their asset holdings, they require risk premiums only for bearing systematic (or non-diversifiable) risk. This systematic or "market" risk, generally measured by "beta" under the CAPM, is the sensitivity of a firm's stock price to market-wide price movements. As measured using APT, systematic risk is measured by the sensitivity of market prices to a number of economic factors, such as changes in real interest rates, unexpected fluctuations in GNP growth, and unanticipated changes in inflation.

Finance theory thus implies that stock market investors are concerned not with the total variability of the firm's cash flows (which we shall refer to hereafter as "total risk"), but only with the *co-variability* of those flows with the performance of the economy as a whole. Finance theorists have therefore maintained that reducing risks at the corporate level which are diversifiable at the portfolio level does not benefit stockholders. Consequently, the argument goes, most company-specific risks, provided they do not significantly raise the prospect of bankruptcy, can be managed more efficiently by stockholders.

Recent scholarship, however, has argued that although total risk may not affect investors' required returns, large unsystematic risks, if unmanaged, can substantially reduce the value of the firm. In terms of the DCF model of firm value, diversifiable risks may not raise investors' discount rates (the denominator), but they can significantly lower the level of the firm's expected cash flows (the numerator). If this is the case, then reducing total risk can increase expected cash flows, thereby increasing the value of the firm. Given these assumptions, which we think quite reasonable, corporate hedging makes economic sense.

How does higher total risk lower expectations about future cash flows? Firms with higher *total* risk, all else equal, are more likely to find themselves in financial distress. Financial difficulties in turn are likely to disrupt the operating side of the business, reducing the level of future operating cash flows. Perhaps most important, financial distress can give rise to management incentives that conflict with the interests of other parties who do business with the firm; and the adverse effect of such incentives on sales and operating costs is compounded by the risk-aversion of customers, managers, employees, suppliers, and other corporate stakeholders. In addition, variability in corporate earnings can affect a firm's ability to take full advantage of tax credits and writeoffs.

Key point

*Financial difficulties are likely to disrupt the operating side of the business, reducing the level of future operating cash flows.*

## The Adverse Incentive Problem

Financial distress, or the threat of bankruptcy, affects management incentives in three fundamental ways. First, managers are more likely to choose high-risk investments that benefit shareholders at the expense of bondholders. Second, they have a tendency to exit promising lines of business or liquidate the entire firm when they would otherwise continue to operate. Third, they may have an incentive to produce goods of inferior quality and provide a less safe work environment for their employees.

A number of finance theorists discuss how the possibility of bankruptcy leads firms to choose possibly suboptimal investment projects that expropriate wealth from their creditors.[2] They demonstrate that if bankruptcy is likely, the firm's stockholders have an incentive to invest in very risky projects, even if they have negative net present values. This is because the bondholders, rather than the stockholders, bear most of the downside risk from these investments, while the stockholders enjoy most of the gain from the upside potential.

Similarly, one of the present authors shows that financial distress, or the threat of bankruptcy, has an important impact on the firm's liquidation decision and, therefore, on its sales and costs.[3] Managers of financially sound firms, as representatives of stockholder interests, generally will not choose to liquidate the firm because stockholders, as the firm's most junior claimants, receive liquidation proceeds only after other claimants are paid in full. Also, because they will lose their jobs if the firm fails, managers have a more direct incentive to keep the firm in business. As the firm progresses through stages of financial distress toward bankruptcy, however, the firm's creditors exert increasing influence on management decisions; and creditors' priority claim to any liquidation proceeds gives them a much stronger incentive to liquidate the firm. The possibility that a financially distressed firm may be liquidated is shown to reduce the current sales and raise the operating costs of high-risk firms by raising concerns of customers, suppliers, and employees.

Because both stockholders and managers have strong incentives to avoid bankruptcy and liquidation, management may take actions under the threat of financial distress they would not otherwise take. For example, a firm having difficulty raising cash may be tempted to lower the quality of its products and services; they may also cut corners on safety for their employees. These temptations will be especially strong in cases where quality or safety is difficult to monitor from the outside, and where the damage may not come to light immediately. Or, a firm facing financial distress may be tempted to conserve cash by cutting back on research and development, advertising and promotional expenditures, and various forms of working capital such as inventory and receivables.

Thus, while a healthy firm has a strong incentive to produce high-quality products and to take other actions that ensure its long-term viability, these normal incentives are likely to change if the firm is suffering financial distress. Under the threat of bankruptcy, the long-run value of a strong reputation may be less important than generating enough cash to make it through the next day. The cost savings associated with cutting quality levels may be particularly attractive to firms facing creditors threatening to take over and possibly liquidate the firm.

Potential customers and other stakeholders anticipate these changes in management incentives and actions. As a result, they become increasingly reluctant to do business with firms in financial distress, as well as with high-risk firms likely to face financial distress in the future. Such expectations of consumers, suppliers, and even employees will adversely affect the firm's future sales, operating costs, and financing costs. In short, the expectation alone of financial distress reduces the expected value of doing business with the firm. key pt

## The Effect on Sales

The incentive of companies in financial distress to produce lower quality products will scare off potential customers. This may, in part, explain Chrysler's difficulty in attracting customers when they were on the verge of bankruptcy. In response to this

2. See, for example, Michael Jensen and William Meckling, "Theory of the -rm: Managerial Behavior, Agency Costs and Ownership Structure," *Journal of -nancial Economics* 3 (1976), pp.305-360; Stewart Myers, "Determinants of -orporate Borrowing," *Journal of Financial Economics* 9 (1977), pp.158-147; and Dan Galai and Ron Masulis: "The Option Pricing Model and the Risk Factor of Stock," *Journal of Financial Economics* 3 (1976), pp.53-81.

3. See Sheridan Titman "The Effect of Capital Structure on a Firm's Liquidation Decision," *Journal of Financial Economics* 13 (1984), pp.137-183.

44 *Reducing total risk can aid a firm's marketing efforts by providing greater assurance to potential customers that the company will be around in the future to service and upgrade its products.*

customer hesitancy, Chrysler decided to offer 5-year service warranties on its new cars. This service contract, which may have been very costly to Chrysler, would probably have been unnecessary if Chrysler were not in such financial straits. Certainly, none of its competitors matched its generous terms.

But product quality is not consumers' only worry; service is a major concern as well. If the original supplier goes out of business, parts and repairs may become a problem. This is particularly important where there are economies of scale in producing parts and providing service. The volume of parts for the aftermarket is likely to be far lower than the volume required for the production of new units. Hence, if the firm goes out of business, production volume will drop precipitously, raising the cost and the price of spare parts. Spare parts will also be more difficult to locate. Similarly, the value of investing in specialized training and equipment to repair a particular product will diminish if the product's sales outlook becomes shaky. This means fewer qualified mechanics and more difficulty in finding those who are around. (The problems roughly parallel those that might befall the owner of a Maserati Biturbo whose car breaks down in an obscure town in the middle of nowhere.)

Thus, reducing total risk can aid a firm's marketing efforts by providing greater assurance to potential customers that the company will be around in the future to service and upgrade its products. Purchasers of long-lived capital assets are especially concerned about the seller's survival. Potential buyers of Chrysler cars, for example, were understandably nervous about purchasing a product which they might have difficulty getting serviced if Chrysler went bankrupt. Clearly, it is comforting to know that the manufacturer will be there to service the equipment and supply new parts as old ones wear out.

At present, with a shakeout under way in the personal computer business, and with several companies dropping by the wayside, shoppers are more worried about producers' staying power. IBM is exploiting buyers' fears about its competitors' longevity with the message, "What most people want from a computer company is a good night's sleep." The widespread concern about whether companies will be around to service and update their machines threatens further to erode the market shares of smaller producers.

In general, whenever customers face large (fixed) costs to switching suppliers, the suppliers have an incentive to present a low-risk image. Switching costs arise when buyers must spend substantial amounts of money tailoring the product to their particular needs, which is often the case in the purchase of a computer or factory automation equipment. These costs come in the form of specific product adaptations, necessary investments in specialized ancillary equipment, and time spent learning how to operate a supplier's equipment or software.

The importance of having a secure position in the marketplace is evident in the case where outside suppliers provide complementary products or services that add value to the firm's product. If a supplier must bear a fixed cost to adapt its product to work with a new brand, it will first make an assessment of the brand's viability in the marketplace. For example, software programs are produced for computers with the largest market share first; only later, if at all, are they rewritten for other brands.

Risk hurts sales for another reason as well. Most firms diversify among suppliers in order to reduce disruption to operations in the event that any one supplier is unable to meet its commitments. It takes time to line up new suppliers, ascertain that their products meet required quality standards, negotiate prices and credit terms, establish new shipping routines, and make all the other adjustments necessary to fit a new supplier into the production schedule. The problem of finding new sources of supply becomes more acute during periods of shortage, such as those caused by price controls. Because financial distress can jeopardize a firm's ability to serve as a stable source of supply, riskier suppliers, all else equal, will gain a smaller share of orders. This is true of suppliers of commodity-type products as well as those of specialized products.

For example, when Wheeling-Pittsburgh Steel Corp. filed for bankruptcy in April 1985, customers reduced their orders, the company was forced to discount prices on some products, and suppliers changed their credit terms to cash-on-delivery. One buyer of Wheeling-Pittsburgh's output interviewed by the *Wall Street Journal* said that he wasn't looking for another supplier because his company had multiple sources. But, he added, "If I were single-sourced, I'd have been on the phone two months ago looking for other suppliers."[4]

4. *The Wall Street Journal*, April 18, 1985, p.4

*Even if a firm facing financial distress intends to be conscientious in dealing with its creditors, it will have difficulty in assuring them of its intentions.*

Risky firms become victims of lost consumer confidence. Customers aren't as willing to do business with a firm that might be going out of business. To some extent, this fear is self-fulfilling as lost sales worsen the firm's financial position and result in a further reduction in sales. Thus, the riskier a firm, the lower its sales are likely to be. Those customers that continue to do business with a risky firm will reduce the price they are willing to pay for the firm's products by an amount equal to their expected damages. As a result, both the unit sales volume and the price received by a firm can be affected by the (perceived as much as the actual) level of its total risk.

As sales decline, distribution will suffer as well, especially for products that cannot just be plugged in, eaten, or worn. Distributors must then invest in training programs so that their salespeople understand and are able to demonstrate the uses of the company's latest products. This fixed cost, in combination with scarce shelf space, limits the number of product lines retailers are willing to carry. The resulting decline in sales further hinders the firm's ability to take advantage of economies of scale, making it less competitive still.

Changes in total risk affect the probability of bankruptcy and, therefore, the possibility that a company will quit a business. Consequently, any action taken by a firm that decreases its total risk will improve its prospects for survival and, hence, its sales outlook.

## The Effect on Operating Costs

A firm's cost of doing business is, in part, a function of its suppliers' view of the company's long-run viability. A firm struggling to survive is unlikely to find suppliers bending over backwards to provide it with specially developed products or services, particularly if those products or services are unique and suitable for use only by the firm in question. In general, the value of investing in a long-term relationship with a customer will depend on whether the customer is expected to survive in the long run. The lower the likelihood of future survival, the more of these relationship costs the customer will have to bear up front in the form of higher prices or less closely-tailored services and products.

Lower-risk firms also have an easier time attracting and retaining good personnel. In the event of liquidation, employees must bear search costs, especially where the job provides firm-specific training and skills not easily transferred elsewhere. Higher level managers bear the stigma of being associated with a failure. The more difficult it is for potential future employers to determine the extent of an individual manager's culpability in the corporate failure, the higher the cost attached to this stigma.

## The Effect on Financing Costs

The potential for myopic behavior on the part of the firm extends to its dealings with creditors as well. A company that expects to remain in business will generally be very protective of its credit reputation. The value of a good credit reputation, however, is lower for firms that may not survive to reap the long-run benefits. Such firms have an incentive to borrow money under false pretenses and mistreat creditors in order to delay the onset of bankruptcy. Because creditors understand this change in incentives, risky firms will find it more difficult to borrow and obtain credit under favorable terms. Moreover, even if a firm facing financial distress intends to be conscientious in dealing with its creditors, it will have difficulty in assuring them of its intentions.

Firms extending trade credit will have a difficult time imposing sufficiently stringent conditions to assure themselves of repayment. Rather than putting up with the associated risks and problems, they are likely just to cut off trade credit for riskier firms. The potential loss of supplier credits presumably is costly since firms seem to prefer it to most other sources of funds, possibly because it is such a flexible form of financing. (The alternative of negotiating bank loans or other financing every time additional credit is needed imposes a variety of transaction costs on the firm that may be avoided with supplier credits.)

Excessive variability in cash flows could also affect the firm's ability to borrow. Just as a firm facing financial distress cannot be trusted to maintain product quality, so it cannot be trusted to honor its obligation to its creditors. Because of the option-like character of equity shares, shareholders have an incentive to select high-risk projects which increase their wealth by reducing the value of the firm's liabilities. As pointed out earlier, equity holders receive *all* of the upside potential from high-risk investments, whereas bondholders share in the downside losses.

Consequently, the riskier a firm is perceived to be, the more stringent the restrictions lenders will

 *The riskier a firm is perceived to be, the more stringent the restrictions lenders will impose on its operating policies and investment projects. These restrictions can prove especially costly for high-growth firms.*

impose on its operating policies and investment projects. These restrictions can prove especially costly for high-growth firms with financing requirements that exceed internally generated cash flow. Moreover, with a risky firm, the interest rate necessary to compensate lenders for the threat that stockholders will yield to temptation may be so high, and the restrictive debt covenants so tightly drawn, as to virtually guarantee that the firm will be unable to pay back its new debts. The result is a drying up of new credits. This could cause the firm to forgo attractive projects, especially if the alternative is an equity issue requiring disclosure of valuable information to competitors.[5]

Decreasing total risk can reduce or eliminate some of the more onerous debt restrictions and covenants. Investment and operating policies with fewer restrictions on them should increase expected future cash flows and shareholder wealth.

## The Problem of Risk Aversion

We have already seen one example of how total risk could change the value of the firm as a going concern. Specifically, creditors might push into bankruptcy and liquidation a firm that would otherwise survive. Recognition of this possibility will be reflected in the form of lower sales and higher operating costs.

Less evident is the likelihood that even in a 100 percent equity-financed firm, the firm's value as a going concern will be reduced by large risk exposures. The income streams of most managers and employees are probably not well (or easily) diversified. Consequently, most people will be concerned about the total risk of their job-derived income. The close connection between corporate and personal risk means that a riskier firm must pay employees more to induce them to commit their human capital to the firm. Similarly, suppliers, distributors, and other corporate-stakeholders will demand to be compensated for bearing added risk. As a result, higher total risk increases the cost of maintaining the organization.

Competition, however, will limit (to the level set by efficient firms) a high-risk company's ability to meet demands for higher compensation. A highly risky firm, therefore, will have difficulty maintaining its organization. The case of AM International is a good illustration of the organizational problems posed by financial distress. After disastrous investments in high-technology businesses, AM International filed for Chapter 11 protection in 1982. The unsettled conditions at the firm, both before and after entering bankruptcy, led to a rapid turnover of managers, with a high cost. As the president of one of AM's divisions commented, "When you are constantly shifting direction, there is no civility, no culture, none of those things that make good companies."[6] The Multigraphics Division had seven presidents in as many years. And 40 executives paraded through the division's six vice-president slots in one four-year period.

The reduction in value caused by such disorder is especially great for those firms whose principal assets are not physical but intangible—assets, for example, that take the form of organizational skills and assets inseparable from the firm itself. One such skill involves knowledge about how best to service a market, including new product development and adaptation, quality control, advertising, distribution, after-sales service, and the general ability to read changing market desires and translate them into salable products. Other valuable organizational assets whose worth would suffer from financial distress—aside from managers with their firm-specific human capital—could include a network of independent distributors or suppliers of specialized products, such as software.

Similarly, many firms are in industries that require salespeople to develop close relationships with their customers. Financial distress, which increases a salesman's personal risk, will cause him—whether he is a high-technology sales engineer or a stockbroker—to jump ship, taking customers with them and further eroding the value of the firm. For example, when Cordis, a medical technology firm, found itself in a battle with the Food and Drug Administration (FDA) over the safety of its pacemakers, it suffered serious damage to its sales force, whose members found their commission income plummeting. Many salespeople switched to Cordis's competitors. Because of the complexity of pacer technology, salespeople tend to have close relationships with the physicians they supply and thus can take their customers along when they

5. This point was made in by Donald Lessard and Alan Shapiro in "Guidelines for Global Financing Choices," *Midland Corporate Finance Journal* (Winter 1983), pp.68-80.

6. *Business Week*, December 3, 1984, p.169.

*The close connection between corporate and personal risk means that a riskier firm must pay employees more to induce them to commit their human capital to the firm.*

switch companies. Needless to say, the cost to the company from desertion amongst their sales force was considerable.

Moreover, even a temporary increase in risk can do permanent damage to the firm if some stakeholders leave in response to their perception of added personal risk. In order to reconstitute the organization, the firm must bear a variety of fixed costs associated with replacing those risk-averse stakeholders. These costs—which include the costs of hiring and training new managers, salespeople and other employees, adding new distributors, and finding new suppliers—reduce the value of the firm as an ongoing entity. When disintegration of the organization has progressed far enough, cash flows will turn negative. At this point, shareholders will have no choice but to liquidate the firm. A creditor-financed firm will likely be liquidated prior to this point.

## Tax Effects

As the variability of operating profits increases, so does the probability that a firm will be unable to make full use of its tax credits and depreciation and interest expense tax deductions.[7] To the extent the resale market in these tax benefits is imperfect (as measured by the discount taken when the tax benefit is sold), an increase in total risk will lead to a reduction in expected corporate cash flows. If the tax credit or tax loss is carried forward, the relevant cost is the reduction in the present value of the tax benefit. By reducing its total risk, a firm can increase the expected value of its tax credits and tax writeoffs and thereby increase its expected future cash flows.

## Why Total Risk Matters: Summing Up

The negative feedback effect of large corporate exposure to risk on expected cash flow should now be evident. As total risk goes up, the firm's cost of doing business rises, reducing its prospects for survival. The combination of risk aversion and poorer corporate prospects weakens the bonds between the firm and the individuals who comprise the extended organization; and an exodus begins. Distributors switch to other brands, suppliers reorient their production facilities, and firms supplying complementary services and specialized products tailor their products for competing brands. Recognizing this, customers buy less; and the firm's best employees either demand higher salaries or leave, taking their firm-specific knowledge elsewhere. This adds to the firm's risk, which, in turn, further affects the firm's sales and cost of doing business. Thus, there is a natural progression from increased total risk to increased risk of bankruptcy and liquidation. Furthermore, even if liquidation is unlikely, total risk will lower the firm's value by an amount equal to the added cost of organizational maintenance.

To summarize the arguments above, the true cost of higher corporate risk is the reduction in the value of the firm's tangible and intangible assets caused by the presence (or probability) of financial distress. In general, the greater the value added by the organization to the firm's products or services and the more expensive it is to reconstitute that organization, the greater this cost is likely to be.

## Characteristics of Firms With High Costs of Financial Risk

Based on the previous discussion, it is possible to identify specific characteristics of firms for which financial distress is especially costly. Such companies would therefore be likely to benefit most from active management of total corporate risk. Some of these characteristics are industry-specific, based on product type, while others are firm-specific. Industry-specific product characteristics include the following:

- *Products that require repairs.* This is illustrated by Lee Iacocca's response to suggestions that Chrysler declare bankruptcy: "Our situation was unique... It wasn't like the cereal business. If Kellogg's were known to be going out of business, nobody would say:'Well, I won't buy their cornflakes today. What if I get stuck with a box of cereal and there's nobody around to service it?'"[8]
- *Goods or services whose quality is an important attribute but is difficult to determine in advance.* One such service is air transportation. In fact, airline

7. This effect was first pointed out by Harry DeAngelo and Ron Masulis in "Optimal Capital Structure Under Corporate and Personal Taxation," *Journal of Financial Economics* 8 (1980), pp.3-29.

8. *Fortune*, November 26, 1984, p.224.

*Firms with substantial growth opportunities will often prefer to maintain a continuous research and development program, and to fund substantial advertising and other marketing expenditures, in both good times and bad.*

companies in financial difficulty have been hurt by the common belief that they are more likely to cut corners on safety, thereby increasing the risk of an accident.

- *Products for which there are switching costs.* Such products would include computers or office and factory automation equipment.
- *Products whose value to customers depends on the services and complementary products supplied by independent companies.* As we saw earlier, many firms require third parties to distribute, sell, service, upgrade, and otherwise add value to their products. Being a low-risk firm helps persuade independent firms to enter into such a symbiotic relationship.

Firm-specific factors include the following:

- *High-growth opportunities.* Firms having more positive net present value projects available than they can finance with internally-generated funds will jeopardize their access to outside financing by the appearance of being risky. Otherwise, prospective investors could be scared off by the previously-discussed management incentive problems.
- *Substantial organizational assets.* Firms whose principal assets are intangible—in the form of managers and employees with firm-specific human capital, outside distributors, suppliers, brand names, a reputation for quality and reliability—will have a higher cost of financial distress than firms with mostly physical assets. These intangible assets will rapidly depreciate in value if the firm experiences, or seems likely to experience, financial distress. As firm risk increases, the value of a reputation for quality products diminishes, and managers and other stakeholders are increasingly likely to sever their ties with the firm.
- *Large excess tax deductions.* Companies such as Chrysler and U.S. Steel cannot take full advantage of their available tax losses, much less the interest on additional debt. Thus, they have less incentive to load up on debt.

## Determining the Firm's Risk Profile

A company's optimal risk profile should be determined by trading off the costs of the firm bearing all (or some) of its risks against the costs of somehow hedging or otherwise reducing those risks. In deciding on an appropriate risk profile, management should conduct a comprehensive analysis of all of its significant exposures. The principal focus of this analysis should be on the risk of *cash insolvency*—that is, the probability of running out of cash before meeting debt servicing charges—given a particular risk profile. Cash insolvency is critical because the inability to meet principal, interest, and lease payments may lead to financial insolvency and, ultimately, to bankruptcy.

However, the analysis can and should be extended to examine the firm's capacity, under various risk scenarios, to service fixed charges of any kind. For example, those firms that perceive large costs to cutting preferred and common stock dividends will treat these as a fixed cost. Strategic factors also enter here. Firms with substantial growth opportunities will often prefer to maintain a continuous research and development program, and to fund substantial advertising and other marketing expenditures, in both good times and bad.

## The Worst Case Scenario

Unfortunately, the difficulty of performing a thorough cash flow analysis may lead firms to limit themselves to using rules of thumb, usually based on various coverage ratios. But coverage ratios do not tell a financial manager what is most important: the probability of cash insolvency associated with alternative risk profiles. This requires a series of cash budgets prepared assuming (1) different economic conditions and (2) the levels of usage of different risk-reducing mechanisms. To do this properly, the financial manager must specify a range of likely future economic scenarios and how the firm's cash flows will be affected by these developments, with a probability attached to each scenario. Moreover, it is necessary to determine other possible sources of cash besides the cash flow from operations. This includes liquid assets that can be drawn down, accounts payable that could be stretched, expenditures that could be deferred, and assets that could be sold. The end result is a series of net cash flows that are or can be generated under each of the different economic scenarios. Based on the associated probabilities, the financial manager can then examine these cash flows and see whether a particular risk profile exposes the firm to too much financial risk.

A useful place to begin determination of an appropriate risk profile is to analyze what happens

to a firm's cash flows under a "worst case scenario." This could mean a general or industry recession when sales are severely depressed, but it could be any combination of adverse circumstances. For example, a company with a sizable export market, such as Rolls Royce, might be most concerned with a possible appreciation in the value of its home currency, while a pharmaceutical company like Johnson & Johnson might be especially concerned with the effects of a product recall.

In 1961, Gordon Donaldson presented a framework for evaluating corporate debt capacity that can be easily adapted (and, in fact, may actually be better suited) to the task of quantifying total corporate risks.[9] In order to use Donaldson's method for this purpose, management would begin by identifying the various sources of risk to which the firm is exposed. Then, for each of these categories of risk, net cash flows can be calculated assuming the worst happens; and, at the end of this process, management will have an estimate of the cash balance the firm can reasonably expect to have at the end of the recession (or some firm-specific catastrophe). Specifically, this means estimating

$$CB_r = CB_o + NCF_r$$

where

$CB_r$ = the cash balance at the end of a recession
$CB_o$ = the cash balance at the start of a recession
$NCF_r$ = the net cash flow during the recession.

By doing this calculation for a range of possible net recessionary cash flows, this information can be used to construct a probability distribution of the ending cash balance, $CB_r$.

The next step in this analysis is to compare these cash flows to fixed charges. Then for each increment of debt, insurance, forward contracts and the like, the firm could determine the probability of cash insolvency based on the probability distribution of $CB_r$.

Suppose, for example, that a firm normally maintains \$1 million in cash and marketable securities. This is the cash balance that would be on hand at the start of a recession or the onset of some other adverse circumstances. Assume that such an economic decline, when it comes, is expected to last for two years. To show the effects of additional debt on the firm's risk profile, assume it borrows an additional \$5 million, with annual debt servicing charges of \$1.5 million. Its cash balance at the end of the economic decline will be

$$CB_r = \$1{,}000{,}000 - (2 \times \$1{,}500{,}000) + NCF_r$$
$$= -\$2{,}000{,}000 + NCF_r$$

Hence, the probability of cash insolvency under this financing plan equals the probability that net cash flow under adverse conditions will fall below \$2 million. Management must then decide how high a probability of cash insolvency it is prepared to tolerate. (This judgment should not be based on its own preferences alone, but rather on what would add most value to shareholders.)

## Cash Inadequacy

Thus far, our analysis of the firm's risk profile has been in terms of the probability of cash insolvency. But long before reaching this point, the firm could be in serious financial trouble. More important, as we saw earlier, this financial distress can prove very costly for certain types of firms, especially those with substantial amounts of organizational assets.

This means that the risk analysis should be extended to deal with the case of *cash inadequacy*, defined by Donaldson as the inability to fund all desired, but not absolutely essential, expenditures.[10] This category would include items such as dividends, an R&D program, expenditures to upgrade plant and equipment, and advertising and other marketing costs. By this point, the company is cutting into muscle and bone; and this will affect its ability to sustain whatever competitive advantage it has. The result will be lower expected future operating cash flows, which will be reflected in a lower stock price today.

Table 1 shows the various ways in which firms can mobilize financial resources in the event of a liquidity problem.[11] These financial resources can be categorized into uncommitted reserves, reductions of planned expenditures, and liquidation of assets. The first category includes excess cash and marketable securities, unused lines of credit, and other sources of liquidity that can be readily

9. *Corporate Debt Capacity* (Boston: Division of Research, Harvard Business School), 1961.

10. Donaldson, *Corporate Debt Capacity* See earlier citation.

11. The mobilization of financial resources is discussed by Donaldson in *Strategy for Financial Mobility* (Homewood, Ill.: Richard D.Irwin), 1971.

TABLE 1
Inventory of Financial Resources

| Resources | Available for use within: One quarter | One year | Three years |
|---|---|---|---|
| **Uncommitted reserves** | | | |
| Instant reserves | | | |
| Surplus cash | $_____ | | |
| Unused line of credit | $_____ | | |
| Negotiable reserves | | | |
| Additional bank loans | $_____ | | |
| Issue of long-term debt | | $_____ | |
| Issue of new equity | | | |
| Preferred stock | | $_____ | |
| Common stock | | $_____ | |
| **Reduction of planned outflows** | | | |
| Volume related | | | |
| Change in production schedule | $_____ | | |
| Scale related | | | |
| Marketing program | | $_____ | |
| R&D budget | | $_____ | |
| Administrative overhead | | $_____ | |
| Capital expenditures | | $_____ | |
| Value related | | | |
| Dividend payments | | $_____ | |
| **Liquidation of assets** | | | |
| Sales of assets | | | |
| Land and real estate | | $_____ | |
| Equipment | | $_____ | |
| Accounts receivable | | $_____ | |
| Inventory | | $_____ | |
| Sales of business units | | | $_____ |
| **Total financial resources available** | $_____ | $_____ | |
| | | $_____ | $_____ |
| | | | $_____ |

Source: Donaldson (1971), "Strategy for Financial Emergencies," *Harvard Business Review* (November–December), p. 72.

accessed. If these sources prove insufficient, the firm can then begin cutting certain non-essential expenditures. Provided the firm is already operating in an efficient manner, these cuts will trade off future cash flows for current cash. Finally, the firm can sell off some of its assets. Again, if the firm was already being run in a lean manner, these asset sales will harm future profitability. The costs associated with these various forms of cash inadequacy and insolvency must be traded off against the costs of reducing risk.

## Methods of Reducing Total Risk

There are many ways in which firms can reduce their total risk, though some are clearly less costly than others. Depending on the methods used, a firm's cost of achieving a given level of risk reduction can vary widely. Management thus has an incentive to select carefully the most cost-effective method of dealing with total risk.

The methods for managing corporate risk can

*A policy of avoiding risky investments ignores the the extent to which a firm can control risk in other ways. After all, companies are in business to take risks, provided such risks promise to be compensated through adequate returns.*

be broken down into two basic categories: "financial" and "real." The principal risk-reducing techniques that can be categorized as financial include lowering the firm's debt-equity ratio, buying or selling forward or futures contracts, and buying insurance. Real adjustments include the adoption of production processes that reduce the degree of operating leverage, avoidance of high-risk projects, and abandonment of existing high-risk products (such as those subject to large liability suits for which no adequate insurance is available).

## Restricting the Debt-Equity Ratio

The finance literature is replete with references to the effect of debt financing on the firm's total risk. By restricting its debt ratio, the firm will reduce its degree of financial leverage, thereby decreasing the probability of financial distress and bankruptcy. In the limit, the all-equity firm would virtually eliminate the probability of bankruptcy (though there would still be a positive probability of voluntary liquidation).[12]

The fact that we see very few all-equity financed firms suggests that this risk-reducing technique is not costless. In particular, because interest payments come out of before-tax income whereas dividends are paid out of after-tax income, debt may be a less expensive source of financing, at least up to a certain point. Hence, the loss of the interest tax shield provided by debt serves as a disincentive to firms to lower their total risk by means of debt reduction. At the same time, however, the tax advantage of debt gives firms an incentive to use other methods of reducing total corporate risk, enabling them to add still more debt to their capital structures. This may allow them to increase their debt tax shield without significantly raising the probability of financial distress.

## Futures and Forward Contracts

A futures or a forward contract calls for delivery, at a fixed future date, of a specified quantity of a given commodity—be it foreign exchange or orange juice—with the price fixed at the time the contract is set. With a forward or futures contract, a firm can lock in its future cost of inputs or sales revenue. For example, an orange juice manufacturer can hedge against the possibility of price fluctuations in its basic raw material by buying orange juice futures. Similarly, a copper mining firm can lock in the revenues from a new mine by selling the output in advance through use of a copper futures contract.

The cost of a forward or futures contract is simply the cost of executing the transaction. In the case of a forward contract, the cost is the spread between the bid and ask price. With a futures contract it is the opportunity cost of the margin amount plus the trading commissions. In a large, active, and well-organized market, such as the foreign exchange market, these costs are likely to be minimal. In other words, the net present value of a futures or forward contract traded in such a market is close to zero. (But, at the same time, the benefits to the corporation from reducing total risk could be substantial.)

## Insurance

Buying insurance is a standard approach to hedging corporate risk. The basic problem with insurance is the large load embedded in its price; that is, insurance rates include a premium over and above the expected losses associated with the policy in order to cover marketing and claims servicing costs, as well as the implicit costs of moral hazard and adverse selection.[13]

Insurance firms do have a comparative advantage in some areas of risk management, such as efficiency in claims service and evaluating safety projects. In addition, the purchase of insurance can act as a signal to debt holders and other claimants that the firm's investment decisions will be geared toward maximizing the value of the firm rather than the value of the firm's equity.[14]

Those firms for which the cost of risk is not that great, based on the characteristics described in Section 2, will choose to self-insure more of their risks. But it should be emphasized that the decision not to insure risks with an insurance company is not the same as choosing to self-insure. The firm has many other means available to reduce risk. It should use

12. There will still be a positive probability of bankruptcy since even all-equity financed firms rely on trade credit.

13. Moral hazard involves the possibility that having an insurance policy will adversely effect (from the insurance company's standpoint) the purchaser's behavior, while adverse selection involves the possibility that the highest-risk firms choose the highest coverage.

14. The comparative advantages of insurance contracts and insurance companies are discussed at length by David Mayers and Clifford Smith in "The Corporate Insurance Decision," *Chase Financial Quarterly* (Spring, 1982), pp.47-65.

*An investment with a negative NPV when evaluated standing alone may have a positive NPV when account is taken of the beneficial effects of risk reduction on the firm's other project cash flows.*

insurance only when the costs of insurance compare favorably with the costs of other risk-reducing techniques.

## Avoiding High Risk Projects

The easiest way to manage risk is to avoid it, which firms can do by screening out high-risk projects. The ease of this approach, however, masks potentially high opportunity costs. The real issue is the degree of risk a company is willing to tolerate and the return required to bear it. A policy of avoiding risky investments ignores the potentially high returns available and the extent to which a firm can control risk in other ways. After all, companies are in business to take risks, provided such risks are recognized, intelligently managed, and promise to be compensated through adequate returns.

In judging the value of undertaking an investment, however, it must be recognized that a project that adds excessive risk to the firm's overall project portfolio may cause financial trouble, and thereby jeopardize all its other activities. An example of this is the case of the Johns Manville Corporation. The high risk associated with its asbestos division resulted in huge product liability suits and increased uncertainty about the financial viability of the firm. Many other firms are very concerned about the huge product liability suits directed against Johns Manville, A. H. Robbins (the Dalkon shield), and other companies. Their response more often than not is to avoid making those products that may subject them to similar liability lawsuits in the future. Again, however, as pointed out above, the opportunity cost of such a policy may be very great.

While some firms may choose to drop overly risky projects or products, there are less drastic steps that can be taken. Since it is the project's contribution to the riskiness of the firm's portfolio of projects—and not simply the risk of the project itself—that matters, one alternative is to choose projects with cash flows that are negatively correlated, thereby hedging or insuring the cash flows of the other projects. Conversely, firms could avoid choosing projects with returns that are highly correlated because of the added likelihood of bankruptcy.

The latter strategy is not costless, however. Investments with differing cash flow patterns are most likely to be those in businesses outside of management's area of expertise, which of course increases the probability that the investments selected will have negative net present values. But, even an investment with a negative NPV when evaluated standing alone may have a positive NPV when account is taken of the beneficial effects of risk reduction on the firm's other project cash flows.

Firms can also reduce the correlation among project returns without venturing into new businesses by diversifying internationally. The relevant issue is whether diversification is the least costly form of risk reduction. For example, firms can also reduce total risk by designing their projects to have lower operating leverage, a subject we turn to now.

## Reducing the Degree of Operating Leverage

Just as reducing the degree of financial leverage lowers total risk, so too does reducing the degree of operating leverage. And for the same reason: it reduces the ratio of fixed to variable costs. For example, if workers can be laid off or fired with relative ease, the more labor-intensive the production process selected, the higher the variable cost-fixed cost ratio, and so the lower the firm's risk. But, if a more capital-intensive production process has lower expected costs, the benefits of risk reduction can come at a high price. Similarly, a firm may forgo the opportunity to take advantage of economies of scale because of the attendant increase in its degree of operating leverage; it may choose instead to build a scaled-down plant that has higher unit costs but involves less risk. In both cases, however, competitors who select the more efficient process will have a cost edge they can use to great advantage.

Long-term sales contracts provide a possible solution to this dilemma. By entering into such a contract, particularly of the "take-or-pay" variety with a minimum floor price, the firm can take advantage of the lower expected cost of the large-scale, capital-intensive process while at the same time reducing its total risk. But if a long-term contract involves simply a transfer of risk to customers, they will demand a price discount for bearing this risk. Under plausible circumstances, however, both the producer and the purchaser can realize risk-reduction benefits from long-term sales contracts. The producer has a guaranteed outlet for its goods while the customer has a stable source of supply.

In some instances, it is not possible to reduce the degree of operating leverage by altering the labor-to-capital ratio. For example, as long as an

*The risk of the firm can be reduced by converting a portion of a worker's income from a strictly contractual claim into a residual or equity claim.*

airplane is in service, it requires a full crew; that is, its ratio of labor to capital is fixed. As another example, labor is a fixed cost for many Japanese firms because they provide lifetime employment for their workers. Thus, changing the labor-to-capital ratio does not change a Japanese firm's degree of operating leverage.

The alternative in these instances is to convert a portion of the worker's income from a strictly contractual claim into a residual or equity claim. This can be done by tying a substantial portion of the worker's expected income to the firm's profitability. Thus, for example several airlines facing financial difficulty have lowered pay levels while offering higher expected bonuses; others have given their employees stock in lieu of higher pay. Similarly, in Japan, most employees receive on the order of one-third their annual income in the form of a year-end bonus tied to the firm's profitability during the year.

## Tradeoffs Among Risk-Reducing Mechanisms

Choosing among the various risk-reducing techniques described in the previous section involves several considerations. We begin this section by offering a few basic principles of risk management:

- First, since the management of total risk entails real costs, the firm should be prepared to pay a positive price to reduce risk.
- Second, the optimal level of risk—the firm's risk-bearing capacity—is found at the point at which the cost of reducing an additional unit of risk is just equal to the benefit from that incremental degree of risk reduction.
- Third, the firm should take advantage of any opportunities to reduce risk at a zero net cost. Needless to say, the firm should seize any risk-reduction bargains; that is, any opportunities to reduce risk at a below-market price. Such bargains are generally to be found, for example, in the form of government-subsidized insurance, such as political risk insurance or export credit insurance.
- Fourth, the firm must bear in mind that real adjustments to reduce risk usually entail real costs, whereas some financial adjustments, such as the use of forward or futures contracts, may be undertaken at a minimal cost.
- Fifth, the firm must take into account the comparative advantages in risk bearing of different institutions. A large multinational industrial firm may have sufficient diversity of operations to self-insure risks ordinarily transferred to an insurance company.
- Sixth, the firm should take into account the effect that reducing one form of risk can have on another form of risk.

The last point, especially, bears some elaboration. Suppose, for example, General Electric sells jet engines to Lufthansa with payment due in one year and set in deutsche marks. GE can hedge its currency risk on this transaction by selling an equivalent amount of deutsche marks forward for dollars for one year. But this does not mean that all risk is eliminated. By transforming a deutsche-mark denominated contract into a U.S. dollar-denominated contract, use of the forward contract is substituting exposure to inflation risk for exposure to currency risk. Without hedging, GE will know how many deutsche marks it will receive in one year, but it won't know the dollar value of those deutsche marks. By hedging, GE will lock in a dollar price for its receivable, but it will not know the purchasing power of those future dollars.

The choice of hedging or not hedging, therefore, depends on which is the bigger risk, inflation risk or currency risk. For most countries with moderate inflation, the answer will surely be currency risk. But for hyperinflationary countries such as Brazil or Mexico, the future purchasing power of the local currency will be less certain than the future purchasing power of the dollar or other strong currency. In this situation, currency risk will be less of a concern than inflation risk.

Similarly, hedging one end of a transaction without hedging the other could result in more risk than not hedging. Suppose Trader Joe buys 4,000 bottles of French champagne to be delivered and paid for in 90 days. The French franc (FF) price is FF100 per bottle which, at the current spot rate of FF1 = $.11, is equivalent to $11 a bottle. If the 90-day forward rate is $.105, Trader Joe can lock in a dollar cost of $10.50 per bottle.

But suppose Trader Joe buys French francs forward to pay for its purchase and the franc depreciates to $.09, while the price of French champagne remains at FF100 per bottle. Trader Joe will now be facing competition from other wine importers whose cost per bottle is only $9.00, $1.50 below its own cost. This competitive pressure will drive down the price at which Trader Joe can sell its champagne.

*Total cash flows of the two firms joined together will exceed the sum of their cash flows operating independently—not because of any synergistic effects but because customers perceive less risk in buying the new product or service.*

Thus, if it hedges its future purchases of champagne, Trader Joe's dollar profit margin will be hurt by a franc depreciation.

Of course, it will benefit from an appreciation of the franc. The important point, though, is that hedging in this case actually increases the variability, and hence the risk, of Trader Joe's profit margin. The reason is that hedging will fix Trader Joe's dollar cost while its dollar price will vary in line with the dollar value of the franc. By not hedging, Trader Joe's dollar cost and dollar revenue will move in unison, thereby preserving a relatively constant dollar margin.

## Reducing the Costs of Risk

An alternative approach to risk management is to reduce the costs associated with risk. There are several ways of doing this, each of which can be used in conjunction with any of the risk-reducing mechanisms discussed previously.

### Merger with a Larger, More Financially Stable Company

A small firm with an innovative product but in a precarious financial position can strengthen its marketing effort by linking up with a larger, less risky company. Potential customers will realize that the company's prospects have improved and worry less about whether its product will be serviced and upgraded in the future. Thus, total cash flows of the two firms joined together will exceed the sum of their cash flows operating independently—not because of any synergistic effects but because customers perceive less risk in buying the new product or service. Additionally, mergers lead to lower total risk because of diversification, which may benefit the acquiring firm's existing marketing efforts.

### Product Compatibility

Another possibility is to produce equipment compatible with the leading manufacturer's product line. Then, even if the original producer becomes financially distressed, there will still be a large support network. Service and spare parts should be readily available and there should be a ready secondary market for the product. This strategy, however, forces companies to compete on the basis of production cost, hurting those firms with a competitive advantage in the design and development of novel products. For example, if Apple Computer had followed the strategy of IBM-compatibility, it would never have brought out its innovative Macintosh personal computer. Also, those so-called "plug compatible" companies (those producing IBM-compatible equipment) have faced enormous risks because their destinies are largely controlled by IBM.

### Off-the-Shelf Components

In line with the previous discussion, the firm can use off-the-shelf components and other readily available product inputs to reduce the cost to its suppliers of financial distress. In the event the firm goes out of business at a later date, the alternative of using specialized inputs could impose heavy costs on suppliers, in the form of plant and equipment and inventory unsuitable for other uses. On the other hand, the decision to stick to off-the-shelf items could prove costly since it limits the firm's design options.

### Training Programs

As pointed out earlier, a major cost of financial distress to employees is that the value of their firm-specific human capital will be greatly diminished should they be forced to seek employment elsewhere. Firms can reduce this cost by providing opportunities to their employees to develop their human potential so it is applicable to a wider variety of circumstances. For example, Procter and Gamble's reputation for providing employees with a post-graduate education in consumer marketing has enabled it to attract top-flight talent. The portability of this knowledge is evidenced by the large numbers of P & G alumni holding top marketing positions in other firms (though such a policy, by making its employees more mobile, is likely to result in increased turnover with all its attendant costs).

### Using Manufacturers' Representatives

Salespeople who specialize in only one product line face a good deal of personal risk in the event something happens to that line. They will demand to be compensated for bearing this risk. An alternative to hiring in-house salespeople is to use manufacturers' representatives who handle a variety of products and lines. This diversification reduces their

personal risk and lowers the amount of compensation they demamd. The cost of using manufacturers' reps, of course, is that the firm's products may not be adequately represented and serviced.

## Summary and Conclusions

The basic message of this paper is that corporate cash flows are influenced by the firm's risk profile. Its revenues, operating costs, financing costs, taxes, and future investment opportunities will all be affected by the likelihood of financial distress, which in turn is a function of total risk. Consequently, even though finance theory maintains that reducing total risk will not lower the firm's required rate of return, it should lead to an increase in corporate cash flows.

For this reason, firms should carefully consider hedging and other risk-reducing activities. Because these activities may be costly, however, it is necessary to balance their costs against the benefits of risk reduction. In particular, the optimal level of risk—the firm's *risk capacity*, if you will—is found at the point at which the cost of reducing an additional unit of risk is just equal to the benefit from that incremental degree of risk reduction. In other words, the firm's risk capacity is defined as the amount of risk it *should* bear, as distinguished from the amount of risk it *could* support.

The risk capacities of firms vary widely given the different natures of the markets they operate in and the strategies they pursue. Some of the characteristics of firms for which financial risk is most costly are industry-specific, based on product type, while others are firm-specific. Industry-specific product characteristics that indicate a high cost of financial distress include the following:

- Products that require periodic repairs.
- Goods or services whose quality is difficult to determine in advance.
- Products for which there are switching costs.
- Products whose value to customers depends on the services and products supplied by independent companies.

Firm-specific factors include the following:

- High-growth companies.
- Substantial amounts of organizational assets.
- Large excess tax deductions.

Firms with high costs of financial distress have relatively low risk capacities and should use the various risk-reducing techniques suggested in this paper. These techniques can be characterized as either "financial" or "real." Financial techniques include limiting the use of debt, buying insurance, and buying or selling forward or futures contracts. Real adjustments include adopting production processes that reduce the degree of operating leverage, avoiding high-risk projects, abandoning existing high-risk products, and choosing projects that lead to internal diversification. Conversely, firms without these characteristics, such as utilities and oil companies, have high risk capacities and, hence, can afford to be more highly leveraged, undertake riskier projects, and self-insure.

In choosing these risk-reducing techniques the firm should take advantage of any opportunities to reduce risk at a zero cost. One consideration that should be borne in mind is that real adjustments to reduce risk usually entail real costs, whereas some financial adjustments, such as the use of forward or futures contracts, may be undertaken at minimal cost. Moreover, the firm should take into account the effect that reducing one form of risk can have on another form of risk. For example, entering into a fixed-price contract eliminates relative price risk but may introduce a significant exposure to inflation risk. It is these kinds of interdependent exposures that an integrated approach for managing corporate risks can recognize and deal with.

## References

DeAngelo, H. and R. Masulis, "Optimal Capital Structure Under Corporate and Personal Taxation," *Journal of Financial Economics* (1980), 8, pp. 3-29.

Donaldson, G., *Corporate Debt Capacity*, Boston: Division of Research, Harvard Business School (1961).

Donaldson, G., *Strategy for Financial Mobility*, Homewood, Ill.:Richard D. Irwin (1971).

Galai, D. and R. Masulis, "The Option Pricing Model and the Risk Factor of Stock," *Journal of Financial Economics* (1976), 3, pp. 53-81.

Jensen, M. and W. Meckling, "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," *Journal of Financial Economics* (1976), 3, pp. 305-360.

Lessard, D. and A. Shapiro, "Guidelines for Global Financing Choices," *Midland Corporate Finance Journal* (Winter 1983), pp. 68-80.

Mayers, D. and C. Smith, "The Corporate Insurance Decision," *Chase Financial Quarterly* (1982), pp. 47-65.

Myers, S., "Determinants of Corporate Borrowing," *Journal of Financial Economics* (1977), 9, pp. 138-147.

Myers, S., "The Search for Optimal Capital Structure," *Midland Corporate Finance Journal* (Fall 1983).

Myers, S., "The Capital Structure Puzzle," *Journal of Finance* (July 1984), pp. 575-592.

Titman, S., "The Effect of Capital Structure on a Firm's Liquidation Decision," *Journal of Financial Economics* (1984), 13, pp. 137-183.

Titman, S. and R. Wessels, "The Determinants of the Capital Structure Choice," UCLA working paper (1984).

**Alan Shapiro**

*is Associate Professor of Finance at UCLA's Graduate School of Management. His book,* Multinational Financial Management *(Allyn and Bacon, 1982) has been adopted by many graduate business schools as their standard international finance text. The focus of Dr. Shapiro's research is international capital budgeting, political and exchange risk management, global financing strategy, and the evaluation of foreign operating performance.*

**Sheridan Titman**

*is an Assistant Professor of Finance at UCLA's Graduate School of Management. He is the author of numerous articles in the area of corporate finance, asset pricing, and real estate finance.*

**David Mayers**
*Ohio State University*

**Clifford W. Smith, Jr.**
*University of Rochester*

# On the Corporate Demand for Insurance: Evidence from the Reinsurance Market*

Significant attention has focused on the determinants of corporate insurance purchases. While this analysis generally involves observable firm characteristics, its implications have been untested. This is primarily due to the difficulty in obtaining data on corporate insurance purchases. We examine one industry where data on insurance purchases are systematically reported: the insurance industry. A reinsurance contract is an insurance policy purchased by one insurance company from another. Our examination of reinsurance purchases by property/casualty insurance companies provides strong evidence on the effects of ownership structure, size, geographic concentration, and line-of-business concentration on the demand for reinsurance.

## I. Introduction

Recent studies have focused on the determinants of the corporate demand for insurance.[1] These analyses explicitly recognize that while the primary motive for individuals' insurance purchases, risk aversion, can partially explain the demand for insurance by closely held corporations and partnerships, it provides a deficient explanation for insurance purchases by widely held corporations. The corporate form is itself a contractual structure with significant risk-management capabilities. Since the corporation's owners, its stockholders, can hold well-diversified portfolios of financial claims, idiosyncratic losses can be managed through diversification. Thus, the analysis has focused on, instead of risk aver-

* Shepard Professor of Insurance, College of Business, The Ohio State University, Columbus, Ohio, 43210, and Clarey Professor of Finance, William E. Simon Graduate School of Business Administration, University of Rochester, Rochester, N.Y., 14627. We thank editors D. Diamond and A. Madansky, an anonymous referee, H. Bessembinder, S. Chaplinsky, H. DeAngelo, S. Harrington, L. Lang, G. Niehaus, M. Smith, and J. Zimmerman for comments and suggestions.

1. See Mayers and Smith (1982, 1987), Main (1983), and Smith and Witt (1985).

(*Journal of Business*, 1990, vol. 63, no. 1, pt. 1)

0021-9398/90/6301-0002$01.50

sion, the structure of the tax code, costs of financial distress including potential investment-incentive effects of a corporation's capital structure, the corporation's ownership structure, comparative advantages in real service production, and the composition of corporate managers' compensation packages.

While this analysis of the corporate demand for insurance generally involves observable firm characteristics, its implications have been untested. This is primarily due to the difficulty in obtaining data on corporate insurance purchases. Under extant accounting rules, corporations are only required to make footnote-level disclosures of insurance purchases if they are material. Yet, there is one industry where data on insurance purchases are systematically reported: the insurance industry. A reinsurance contract is an insurance policy purchased by one insurance company, the ceding company, from another, the reinsurer. Hence, within the insurance industry, reinsurance purchases are like traditional insurance purchases by industrial corporations.

In this article, we empirically examine the determinants of reinsurance purchases for a sample of 1,276 property/casualty insurance companies. These data include firms across a broad range of ownership structures—stocks, mutuals, Lloyd's, reciprocals. Moreover, we distinguish among stocks that are widely held, closely held, owned by a single family, owned by a mutual, and owned by an association. For some purposes we avoid the complex ownership structure problems implied by subsidiaries and group membership by focusing on a subset of 330 nonsubsidiary nongroup firms.

In Section II, we review the basic hypotheses about the corporate demand for insurance. We describe our data in Section III, analyze the evidence from reinsurance purchases in Section IV, and present our conclusions in Section V.

## II. The Corporate Demand for Insurance

Insurance purchases affect the firm's current market value through changing tax liabilities, contracting costs, or incentives with respect to real investment decisions for either the corporation or its claimholders. Although each provides a potential motive for corporate insurance purchases, they are not mutually exclusive. In the rest of this section we discuss the determinants of corporate insurance purchases and indicate specific adaptations for reinsurance purchases by insurance companies.

*Taxes.* The tax code provides incentives for firms to purchase insurance.[2] The provisions of the code imply a convex tax function for

2. For a more complete analysis of this incentive see Mayers and Smith (1982), Smith and Stulz (1985), or Smith and Witt (1985).



low levels of taxable income and an essentially linear function for taxable income above $100,000.[3] The convexity implies corporations have expected tax liabilities greater than the tax liabilities associated with their expected pretax income. Therefore, the corporate demand for insurance will be generally greater for firms with expected income in the convex region of the tax schedule or with more volatile pretax income.

Insurance companies' ability to deduct indemnity payments to policyholders while investing in tax-exempt municipal bonds makes the convex section of the tax function more important than for manufacturing firms of similar size (see Hendershott and Koch 1980). Since insurance firms typically face a significant probability of taxable income within the convex region, the purchase of reinsurance can reduce the firm's expected tax liability by reducing the volatility of pretax income. A second tax incentive is provided for insurance companies that are members of groups. Reinsurance can transfer profits within the group, allowing recognition of profits so that group taxes are reduced.[4] Thus, we expect to observe more reinsurance by group members than by similar unaffiliated insurance companies.

*Expected bankruptcy costs.* Transactions costs of bankruptcy can induce corporations to purchase insurance since the probability of incurring the costs is reduced by shifting risk to the insurance company. Warner's (1977) evidence that direct bankruptcy costs are less than proportional to firm size suggests small corporations are more likely to purchase insurance. Expected bankruptcy costs should also be more important for firms with higher cash-flow volatilities. For insurance companies, group membership can reduce the demand for externally provided reinsurance by providing a substitute mechanism for lowering expected bankrupty costs through pooling.

The probability of bankruptcy had additional importance for insurance companies: product quality is a negative function of the firm's default risk. Insurance purchasers assess the probability of default and adjust their demand prices accordingly. Moreover, since risk-averse purchasers pay a premium over the actuarially fair rate to eliminate a risk, an insurance company has incentives to reduce its default probability with reinsurance.[5]

*Investment incentives.* Myers (1977) shows that firms have incen-

3. The introduction of the alternative minimum tax in 1986 imposes potentially important nonlinearities above $100,000. However, our data are from 1981, prior to the introduction of this provision.

4. Note that a significant component of state taxation of insurance firms is through a premium tax. To the extent that premium income is taxed, rather than profits, this argument for reinsurance is reduced.

5. See Johnson and Stulz (1987). Fixed costs reflected in insurance premiums are likely sufficient to assure that reinsurance dominates the alternative solution of the insured diversifying across insurance companies.

74

tives to forgo valuable investment opportunities. Myers argues that in some circumstances, with risky debt in the capital structure, taking a positive net present value project makes stockholders worse off because the project's benefits accrue to the bondholders. Mayers and Smith (1987) demonstrate that in certain cases, the purchase of insurance controls this underinvestment incentive.

Insurance companies also can have underinvestment problems. For example, consider a company that experiences an abnormally large loss that reduces the value of both the equity and the firm's outstanding policies. The equity holders now might rationally choose to reject a positive net present value project because the benefits accrue primarily to the policyholders who have prior claim on the firm's assets. However, if the firm had purchased reinsurance, the loss would be indemnified, and the incentive to forgo the value-increasing project would be reduced. These problems are expected to be more severe the smaller the firm's capitalization and the more volatile its cash flows.[6]

Reinsurance also facilitates intrafirm specialization in investment management. For firms with subsidiaries, asset control is maintained by the parent company while regulatory requirements are met by the subsidiary through reinsurance with the parent company.[7] Thus, the assets ultimately backing the policy sold by a subsidiary appear on the parent's, not the subsidiary's, balance sheet.

*Optimal risk sharing.* Closely held corporations are more likely to purchase insurance than firms with less concentrated ownership for the same reason that individuals purchase insurance: risk aversion. Insurance contracts allow owners of closely held firms to specialize in risk-bearing only in dimensions in which they have expertise and thus a comparative advantage (see Arrow 1974, ch. 5).

Ownership structure varies within the insurance industry. For example, there are Lloyd's associations, where insurance contracts are offered by individual underwriters; stock companies that employ the standard corporate form; and mutuals and reciprocals that are more like cooperatives with customer and owner functions merged. Stock

6. While we do not employ data on company capitalization, given the different accounting conventions used across ownership structures, one must be careful when empirically testing this proposition.

7. Indirect evidence of this use of reinsurance by subsidiaries is provided in Mayers and Smith (1989). We find that parent chief executive officer (CEO) compensation is more closely related to assets as a measure of firm size, while subsidiary CEO compensation is more closely related to net premiums written, a sales measure. Additionally, Best (1982) indicates that "100% reinsured" with the parent, where the parent is also a property/casualty company, is common. Given the other control mechanisms available between parent and subsidiary firms, full reinsurance coverage is feasible. However, a reinsurance contract with an unaffiliated reinsurer is likely to restrict coverage by specifying deductibles, coinsurance provisions, and upper limits to control incentive problems. See Huberman, Mayers, and Smith (1983).



company ownership structure also varies from single-owner companies to those that are widely held. Because of variation in risk exposure, ownership structure can be an important determinant of cross-sectional differences in reinsurance purchases: closely held stocks and Lloyd's are expected to reinsure more than firms with less concentrated ownership.

*Real-service efficiencies.* Insurance firms develop a comparative advantage in processing claims because of scale economies and gains from specialization. Thus, noninsurance corporations can increase expected net operating cash flows by purchasing insurance when insurance companies are the low-cost supplier of these services.

Reinsurance firms regularly provide a set of services to ceding insurance companies. The reinsurer frequently has broader experience with low probability events and provides information on pricing and claims adjustment services in particular areas.[8] This information is more likely to be valued highly by small insurance firms, especially those that are geographically diversified or that offer insurance across many lines. Therefore, real-service efficiency arguments can explain both reinsurance purchases by insurance companies in addition to insurance purchases by nonfinancial corporations.

Reinsurance can also be a specialized form of financing. Reinsurance reduces insurance in force, thus relaxing the regulatory constraint on the ratio of capital to insurance in force. This motive is likely to be especially important for mutuals since they cannot raise capital by selling equity.

## III. Data Description

Our basic data are from the A. M. Best Company (Oldwick, New Jersey); their 1981 line-of-business file (Best's) contains data on premiums, losses, and expenses categorized into 26 insurance lines for a large sample of property/casualty insurance firms. The file also identifies each firm's ownership structure (Lloyd's, stock, mutual, or reciprocal), group membership, and reports total admitted assets and the number of states licensed. The file contains usable data on 1,276 firms: 854 stock companies, 320 mutual companies, 60 reciprocal associations, and 42 Lloyd's associations.

Our measure of reinsurance activity is the ratio of reinsurance premiums ceded to total business premiums. Total business is defined as direct business plus reinsurance assumed. Direct business is gross premiums (including policy and membership fees written and renewed

8. See, e.g., Bickelhaupt (1983, p. 824): "Reinsurers also offer many technical advisory services to new insurers or those expanding to new types of insurance or territories."

TABLE 1 Summary Statistics on Distribution of Ratio of Reinsurance Premiums Ceded to Total Business Premiums for 1,276 Property/Casualty Insurance Companies[a]

| Moments | | Quantiles | |
|---|---|---|---|
| Mean | .38 | 100% MAX | 1.33 |
| SD | .33 | 99% | 1.00 |
| Skewness | .63 | 95% | 1.00 |
| Kurtosis | −.80 | 90% | .99 |
| | | 75% Q3 | .62 |
| | | 50% MED | .31 |
| | | 25% Q1 | .10 |
| | | 10% | .02 |
| | | 5% | .00 |
| | | 1% | .00 |
| | | 0% MIN | −.29 |

[a] Total business premiums are defined as direct business written plus reinsurance assumed.

during the year) less return premiums. Reinsurance assumed is the premium income from supplying reinsurance services.[9]

Table 1 displays summary information on the distribution of the ratio of reinsurance premiums to total business premiums for the 1,276 property/casualty insurance companies. Note that the ratio is not bounded by zero and one; the minimum in our sample is −0.29 and the maximum is 1.33. This occurs primarily because of temporal mismatches in income flows. For example, negative values for reinsurance ceded can result from a return of premium by the reinsurer. Conversely, the ratio can exceed one for a firm that has decided to exit from a line of business or a state because it has stopped issuing new policies but reinsures policies in force. We do not believe that this temporal mismatching represents a significant problem for our analysis since there is no apparent reason for it to introduce bias in our procedures. Moreover, the total number of observations of this ratio that lie outside the zero-to-one range is less than 2% of the 1,276 observations.[10]

Our discussion suggests several proxy variables to explain the cross-sectional variation in the ratio of reinsurance premiums ceded to total business premiums.

*Size.* Firm size affects the demand for insurance through taxes, expected bankruptcy costs, investment incentives, and real-service efficiencies. The real-service and bankruptcy-cost effects are straightforward: larger firms should have a lower demand for reinsurance for these reasons. Expected bankruptcy costs are more likely to be an important factor for small firms, and small firms are less likely to have the specialized internal talent available in larger firms. However, the effects of size through the tax and investment-incentive motives are ambiguous.

We measure firm size by total admitted assets. Means and medians for admitted assets by ownership classification are reported in columns 4 and 5 of table 2.

*Business concentration.* We expect business concentration to be closely related to the real-service benefits of reinsurance. Other things equal, the less concentrated the insurer's business across lines of insurance, the more valuable the reinsurer's rating information. However, to make specific predictions about the impact through taxes, expected bankruptcy costs, and real-investment incentives requires additional knowledge of the expected cash-flow volatilities for specific lines as well as potential exposure across lines to common underlying factors such as liability rule changes. For example, firms with high line-of-business concentration could specialize in low-volatility lines of business. The impact of business concentration on the demand for reinsurance is thus ambiguous.

We measure line-of-business concentration by the Herfindahl index of concentration across lines of business.[11] Means and medians for our line-of-business concentration measure by ownership classifications are reported in table 2, columns 6 and 7.

*Geographic concentration.* Geographic concentration can affect reinsurance purchases for three reasons: (1) It increases the volatility of taxable income and thus increases tax-related incentives to reinsure. (2) It increases the volatility of firm value and thus increases incentives to reinsure because of expected bankrupty costs and investment incentives. (3) It reduces the value of real services provided by the insurer and hence reduces the demand for reinsurance. Thus, whereas tax, expected bankrupty costs, and investment incentives all imply a positive association between reinsurance activity and geographic concentration, the real-service incentive implies a negative association.

We measure geographic concentration by the negative of the number of states licensed. Means and medians for the number of states licensed by ownership classification are reported in table 2, columns 8 and 9.

9. We considered other measures of reinsurance activity, specifically measures of net reinsurance activity as opposed to just reinsurance purchased. However, for our purposes, it seems inappropriate to treat reinsurance assumed as simply the negative of reinsurance purchased (ceded). For example, one component of the reinsurance market operates like a simple pooling contract. Under such a contract the average pool member's net reinsurance activity would be zero, just like a firm with no participation in the reinsurance market.

10. Observations outside the zero-to-one range appear symmetrical with 13 negative and 12 greater than one.

11. The Herfindahl index is calculated for each company as

$$H = \Sigma S_L^2,$$

where $L$ stands for line of insurance and $S_L = PI_L/TPI$; $PI_L$ is the dollar amount of direct business written in a particular line of insurance and $TPI$ is the dollar amount of direct business totaled across all 26 lines of insurance.

TABLE 2 Mean (Median) Admitted Assets, Business Concentration, Geographic Concentration, and Best's Rating for Total Sample of 1,276 Firms and Nonsubsidiary/Nongroup Subsample of 330 Firms in the Property/Casualty Insurance Industry

| | Ownership Classification Totals | | | Admitted Assets (in Thousands) | | Business Concentration (Herfindahl) | | Geographic Concentration (No. of States) | | Best's Rating |
|---|---|---|---|---|---|---|---|---|---|---|
| Ownership Classification | Total Company Type (1) | Group Members (2) | Nonsubsidiary/ Nongroup (3) | Total Sample (4) | Nonsubsidiary/ Nongroup (5) | Total Sample (6) | Nonsubsidiary/ Nongroup (7) | Total Sample (8) | Nonsubsidiary/ Nongroup (9) | Nonsubsidiary/ Nongroup (10) |
| 1. Lloyd's associations | 42 | 35 | 7 | 2,472.6 (745.3) | 3,033.4 (2,234.7) | .52 (.49) | .36 (.35) | 1.5 (1.0) | 3.4 (1.0) | 2.0 (0) |
| 2. Mutuals | 320 | 114 | 206 | 154,567.5 (19,794.5) | 41,554.5 (13,249.3) | .43 (.34) | .46 (.38) | 10.9 (3.0) | 6.7 (2.0) | 4.8 (6) |
| 3. Reciprocal associations | 60 | 27 | 33 | 172,535.4 (13,571.0) | 85,991.5 (9,740.9) | .63 (.53) | .71 (.64) | 10.3 (2.0) | 3.4 (1.0) | 2.2 (0) |
| 4. Mutual-owned stocks | 118 | 110 | ... | 110,636.4 (16,372.1) | ... | .42 (.37) | ... | 15.4 (6.5) | ... | ... |
| 5. Association-owned stocks | 38 | 12 | 20 | 90,326.3 (22,954.9) | 40,707.8 (18,492.0) | .58 (.51) | .66 (.71) | 9.3 (3.0) | 9.5 (2.5) | 2.4 (0) |
| 6. Single-owner stocks (subsidiaries) | 80 | 38 | ... | 41,702.9 (8,536.9) | ... | .51 (.44) | ... | 8.4 (2.0) | ... | ... |
| 7. Closely held stocks (subsidiaries) | 36 | 16 | ... | 17,677.9 (5,591.5) | ... | .60 (.52) | ... | 6.7 (2.0) | ... | ... |
| 8. Widely held stocks (subsidiaries) | 500 | 436 | ... | 259,474.1 (31,926.9) | ... | .41 (.32) | ... | 23.6 (16.5) | ... | ... |
| 9. Single-owner stocks (nonsubsidiaries) | 22 | 1 | 21 | 6,996.4 (3,677.8) | 6,592.6 (3,607.7) | .68 (.65) | .66 (.55) | 4.0 (1.0) | 4.2 (1.0) | 2.5 (3) |
| 10. Closely held Stocks (nonsubsidiaries) | 32 | 6 | 26 | 16,616.1 (7,781.6) | 11,047.9 (6,488.5) | .65 (.57) | .71 (.62) | 8.8 (2.0) | 5.6 (1.0) | 2.1 (1) |
| 11. Widely held Stocks (nonsubsidiaries) | 28 | 11 | 17 | 321,291.5 (40,649.6) | 62,774.8 (19,226.9 | .57 (.52) | .68 (.64) | 21.1 (15.5) | 15.7 (13.0) | 3.2 (5) |

NOTE.—Numbers in parentheses are medians.



*Ownership structure.* Although ownership structure is ultimately endogenous, from an econometric perspective, it is a predetermined variable with respect to reinsurance purchases. We control for differences in ownership structure with dummy variables representing different ownership classifications. Best's classifies firms as Lloyd's, stocks, mutuals, and reciprocals. We augment Best's classifications by classifying stock companies as ultimately owned by an association,[12] by a single family (at least 50% owned by one family), as closely held (100 or fewer shareholders), or as widely held (more than 100 shareholders). We also indicate whether the stock company is a subsidiary and for all companies whether they are a member of a group.[13]

Our hypotheses explaining variation in reinsurance purchases across ownership structures imply that the more significant the fraction of total wealth that the insurance company represents to its owners, the greater will be the demand for reinsurance services by the insurance company.[14] Thus we expect the Lloyd's to have the greatest demand for reinsurance followed in order by single-family, closely held, and widely held stocks.[15] We also expect subsidiaries and group members to reinsure more. However, the relative ordering of mutuals, reciprocals, and stocks owned by associations depends on the relative importance of the financing aspects of reinsurance, the potential reduction in expected bankruptcy costs from the issuance of assessable policies (reciprocals), and the specific knowledge of risks for insurance offered to association members.

Groups and subsidiaries pose potential problems for our analysis. In a group the interfirm affiliation can reflect an ownership relation (e.g., a parent company with several subsidiaries could constitute a group) as well as other relations. Companies in groups frequently have "group business pooling" arrangements with the other members of the group. To the extent that these complex contractual arrangements are not captured by our dummy variables, including group members introduces a potential bias in our estimated coefficients. There is also the potential for bias if, for example, firms in groups tend to concentrate in different lines of business than nongroup firms. Reinsurance contracts

12. Examples of associations that own stock insurance companies are the American Medical Association, California Farm Bureau Federation, Blue Shield Association, Catholic Mutual Relief Society of America, and AGWAY, Inc.—a farm supply and food marketing cooperative.

13. This additional information (except for group membership, which is on the file) is obtained from the 1982 *Best's Insurance Reports*.

14. Hence, we assume that variation in individual risk aversion is not sufficient to allow sorting by individuals to wash out cross-sectional variation in reinsurance demand.

15. That some Lloyd's have no specific limitation on the underwriters' liability should be sufficient to generate a difference between Lloyd's and single family stocks. Three of the seven Lloyd's that are not members of insurance groups indicate limitations on underwriter's liability in the 1982 *Best's Insurance Reports*.

between parent and subsidiaries can bias our measure of firm size for subsidiaries and cause problems in interpreting the relation between firm size and reinsurance. For subsidiaries, the causality could be reversed—subsidiaries have smaller measured admitted assets because they reinsure. While one solution to the problem of groups and subsidiaries would be to treat related firms as entities, that consolidation poses problems for at least one reason: the Best's sample is not an exhaustive listing of all insurance companies, hence we would end up measuring partial entities.

We control these firm-definition problems by restricting part of our analysis to that subset of firms where these problems are less severe: the nonsubsidiary, nongroup property/casualty insurance companies. Although this restriction reduces our sample of companies by 74% (as indicated in table 2), it allows more focused tests of our hypotheses. For example, our ownership structure definitions for stand-alone insurance companies are likely more closely related to our hypotheses than they are for subsidiaries: a subsidiary that is ultimately closely held is more likely a part of a larger portfolio of companies. Moreover, including subsidiaries adds dimensions to the ownership structure problem that we do not fully understand. For example, small property/casualty insurance companies that are ultimately closely held are frequently separated from their owner(s) by a holding company whose only asset is the subsidiary insurance company.[16]

*Best's rating*. The default risk of the insurer affects the demand for reinsurance through investment incentives and expected bankruptcy costs. Both arguments imply that riskier firms have incentives to purchase more reinsurance. As a proxy for default risk we employ Best's General Rating from the 1982 *Best's Insurance Reports—Property Casualty*. Best's assigns a group rating where companies operate under an intercompany reinsurance arrangement or where common management and underwriting prevail without a business pooling arrangement. Thus Best's ratings are easiest to interpret for our subset of 330 nonsubsidiary, nongroup firms.

Best's ratings range from A+ to C, but for 85 firms of our sample of 330 (25.8%) a rating is not assigned. Best's lists several reasons for nonassignment, for example, not qualifying for the minimum rating of C and having less than 5 years of continuous operating experience. We assign a value of six for companies with a Best's rating of A+, five for A, four for B+, down to one for C. We assign a value of zero for those firms with no rating. In table 2, we report the average Best's rating for each ownership structure class.[17]

16. We believe that one reason for this organizational structure is that it allows the firm's owners to reduce the regulatory constraints on leverage by allowing debt issuance by the holding company.

17. The qualitative interpretations of our analysis are unchanged if we use a separate indicator variable for each Best's rating rather than these numerical assignments.



*Lines of business*. We control for variation in the demand for reinsurance across lines by using each company's percentage of direct business in each of 23 lines as control variables.[18] Table 3 reports means and ranges for the percentages of direct business by line and ownership classification. The average Lloyd's in our sample does almost 41% of its direct business in commercial multiple peril and the average mutual does about 14% of its direct business in auto physical damage.[19] Hence, there appears to be considerable line specialization by ownership class.

In our analysis of the corporate demand for insurance we have argued that it is important to hold things like the lines in which the firm operates constant in order to examine the impact of size, business concentration, geographic concentration, Best's rating, and ownership structure. However, we suspect that the line-of-business data reported by Best's are not ideal for our purposes. One problem with the data is aggregation: Best's data treat all fire insurance policies, for example, as having identical risk factors. This poses potential problems in the interpretation of our results that we discuss in more detail below.

## IV. Evidence from the Reinsurance Market

*Ownership structure*. We first examine whether, consistent with our hypotheses, the ratio of reinsurance ceded to total business varies across ownership structures. These tests have low power because they do not control for potentially important variation in company size, business concentration, geographic concentration, Best's rating, or lines of business. However, because the tests are simple pairwise comparisons, they allow us to use nonparametric as well as parametric statistics. Consistency between the parametric and nonparametric results strengthens our confidence about the parametric regression results that follow, where we control for the other factors.

Table 4 contains three regressions of reinsurance activity on ownership class dummy variables (intercepts are omitted to avoid singularity). Regression 1 employs the total sample of 1,276 companies to estimate the mean reinsurance ratio for each ownership class. Regression 2 also employs the full 1,276 insurance company sample but includes additional dummy variables indicating group membership and interactions between group membership and ownership classification.

18. One of the 26 lines of business reported by Best's (miscellaneous) was originally omitted to avoid singularity in the regression matrix for analysis on the entire sample of insurance companies (1,276). Since none of the 330 firms in the final sample (which omits subsidiaries and group members) write business in reinsurance or international, those two lines are also omitted in the analysis of that sample.

19. Regulation is likely to account for some of the observed line-of-business concentration in Lloyd's. The state laws authorizing American Lloyd's frequently provide that only certain types of insurance may be written by Lloyd's groups.

**TABLE 3** Summary Statistics (Mean/Range) for Percentages of Direct Business Written in Each of 23 Lines of Insurance by Ownership Classification for 1,276 Property/Casualty Insurance Companies

| Line of Insurance | Lloyd's Associations | Mutuals | Reciprocal Associations | Mutual-Owned Stocks | Association-Owned Stocks | Single-Owner Stocks (Subsidiaries) | Closely Held Stocks (Subsidiaries) | Widely Held Stocks (Subsidiaries) | Single-Owner Stocks (Nonsubsidiaries) | Closely Held Stocks (Nonsubsidiaries) | Widely Held Stocks (Nonsubsidiaries) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Fire | .1182 (.667) | .1085 (1.000) | .0782 (.786) | .0491 (.788) | .0175 (.200) | .0530 (.764) | .0312 (.323) | .0502 (.883) | .0332 (.512) | .0730 (1.000) | .0814 (1.000) |
| Allied lines | .0936 (.390) | .0528 (1.000) | .0374 (.437) | .0299 (.640) | .0420 (1.000) | .0303 (.996) | .0390 (1.000) | .0202 (.333) | .0091 (.086) | .0130 (.237) | .0129 (.113) |
| Farmowners multiple peril | .0013 (.020) | .0470 (.646) | .0014 (.051) | .0064 (.371) | .0231 (.169) | .0004 (.012) | .0031 (.059) | .0035 (.122) | .0005 (.012) | .0005 (.014) | .0047 (.121) |
| Homeowners multiple peril | .0749 (.607) | .2620 (.0955) | .0396 (.821) | .1936 (1.000) | .0824 (.780) | .0406 (.625) | .0967 (.976) | .0801 (.856) | .0983 (.858) | .0307 (.562) | .0735 (.429) |
| Commercial multiple peril | 4090 (.965) | .0642 (.926) | .0652 (.864) | .0790 (.863) | .0567 (.928) | .0521 (.450) | .0095 (.150) | .0718 (.885) | .0323 (.645) | .0222 (.274) | .0233 (.233) |
| Ocean marine | .0000 (.000) | .0061 (1.000) | .0168 (1.000) | .0089 (.750) | .0049 (.180) | .0065 (.171) | .0335 (1.000) | .0090 (1.000) | .0000 (.000) | .0014 (.046) | .0006 (.008) |
| Inland marine | .0728 (.827) | .0228 (.934) | .0448 (.8511) | .0223 (.432) | .0065 (.055) | .0172 (.226) | .0171 (.124) | .0484 (1.000) | .0157 (.267) | .0072 (.065) | .0554 (.963) |
| Medical malpractice | .0008 (.016) | .0585 (1.000) | .2031 (1.000) | .0069 (.606) | .2368 (1.000) | .0133 (3.73) | .0007 (.027) | .0128 (1.000) | .0403 (.886) | .0036 (.062) | .1328 (1.000) |
| Earthquake | .0000 (.000) | .0002 (.018) | .0005 (.023) | .0003 (.011) | .0000 (.000) | .0001 (.003) | .0000 (.000) | .0011 (.093) | .0009 (.019) | .0016 (.050) | .0001 (.001) |
| Group accident and health | .0077 (.150) | .0058 (.879) | .0071 (.231) | .0075 (.449) | .0445 (.999) | .0024 (.139) | .0000 (.000) | .0107 (1.000) | .0000 (.000) | .0000 (.000) | .0184 (.314) |
| Credit accident and health | .0000 (.000) | .0000 (.001) | .0000 (1.000) | .0000 (.000) | .0000 (.000) | .0000 (.000) | .0108 (.391) | .0029 (1.000) | .0194 (.427) | .0000 (.000) | .0000 (.000) |
| Other accident and health | .0025 (.102) | .0050 (1.000) | .0001 (.004) | .0115 (.550) | .0003 (.018) | .0120 (.837) | .0006 (.017) | .0088 (1.000) | .0000 (.000) | .0013 (.029) | .0796 (1.000) |
| Workmens' compensation | .0099 (.282) | .0553 (1.000) | .0756 (1.000) | .0786 (.973) | .0727 (.857) | .0657 (1.000) | .1588 (1.000) | .1488 (1.000) | .0475 (1.000) | .1041 (1.000) | .0480 (.355) |
| Other liability | .0215 (.274) | .0372 (1.000) | .0626 (1.000) | .0528 (1.000) | .0323 (.332) | .1400 (1.000) | .0974 (1.000) | .1163 (1.000) | .0361 (.569) | .0596 (.565) | .0734 (1.000) |
| Auto liability | .0421 (.381) | .1448 (.733) | .1881 (.939) | .2588 (.799) | .2290 (1.000) | .2038 (.894) | .2400 (.935) | .1789 (.840) | .1742 (.674) | .2699 (.928) | .2130 (.803) |
| Auto physical damage | .1194 (.999) | .1074 (.625) | .1615 (1.000) | .1663 (.516) | .1290 (.500) | .2181 (1.000) | .1436 (.908) | .1622 (.995) | .2871 (1.000) | .2409 (1.000) | .1033 (.400) |
| Aircraft | .0000 (.000) | .0004 (.103) | .0162 (.973) | .0004 (.038) | .0000 (.000) | .0018 (.109) | .0195 (.546) | .0070 (1.000) | .0000 (.000) | .0329 (1.000) | .0009 (.026) |
| Fidelity | .0003 (.004) | .0003 (.020) | .0001 (.003) | .0039 (.401) | .0001 (.004) | .0037 (.166) | .0000 (.001) | .0048 (1.000) | .0040 (.082) | .0116 (.333) | .0018 (.019) |
| Surety | .0005 (.011) | .0064 (.990) | .0003 (.018) | .0044 (.377) | .0216 (.820) | .0745 (1.000) | .0919 (1.000) | .0158 (1.000) | .1540 (1.000) | .1253 (1.000) | .0322 (.577) |
| Glass | .0001 (.003) | .0002 (.003) | .0003 (.014) | .0002 (.005) | .0002 (.005) | .0001 (.002) | .0000 (.001) | .0003 (.010) | .0000 (.001) | .0001 (.001) | .0002 (.002) |
| Theft | .0015 (.039) | .0009 (.067) | .0003 (.0007) | .0007 (.015) | .0002 (.0007) | .0005 (.008) | .0001 (.002) | .0019 (.326) | .0000 (.000) | .0001 (.001) | .0013 (.020) |
| Boiler | .0001 (.002) | .0033 (.245) | .0000 (.000) | .0028 (.212) | .0000 (.000) | .0000 (.001) | .0000 (.000) | .0014 (.176) | .0000 (.000) | .0000 (.000) | .0357 (.998) |
| Credit | .0049 (.205) | .0008 (.002) | .0000 (.000) | .0000 (.001) | .0000 (.000) | .0000 (.002) | .0000 (.000) | .0113 (1.000) | .0017 (.037) | .0000 (.000) | .0000 (.000) |

NOTE.—The first number in each pair expresses the mean; the second number—in parentheses—expresses the range.

TABLE 4 Cross-Sectional Regressions of Reinsurance Activity (Ratio of Reinsurance Premiums Ceded to Total Business Premiums) on Ownership Classification Dummy Variables for Total Sample of 1,276 Firms and the Nonsubsidiary, Nongroup Subsample of 330 Firms in the Property/Casualty Insurance Industry*

| | Variables | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Regression No. | Lloyd's Assoc. | Mutual | Recip Assoc. | Assoc. Owned | Mutual Owned (Sub.) | Single Owner (Sub.) | Closely Held (Sub.) | Widely Held (Sub.) | Single Owner (Nonsub.) | Closely Held (Nonsub.) | Widely Held (Nonsub.) | Group | $N$ |
| 1 | .88 | .25 | .26 | .29 | .43 | .33 | .52 | .46 | .25 | .27 | .18 | ... | 1,276 |
| | (19.32) | (15.03) | (6.93) | (5.97) | (15.75) | (10.05) | (10.56) | (34.90) | (3.91) | (5.24) | (3.18) | | |
| 2 | .79 | .23 | .23 | .24 | .28 | .30 | .42 | .23 | .26 | .22 | .15 | .27 | 1,276 |
| | (7.22) | (11.44) | (4.54) | (4.23) | (2.78) | (6.76) | (6.52) | (6.23) | (4.07) | (3.90) | (2.08) | (7.00) | |
| † | −.16 | −.22 | −.19 | −.12 | −.12 | −.21 | −.05 | ... | −.49) | .01 | −.19 | ... | ... |
| | (−1.27) | (−4.29) | (−2.26) | (−1.14) | (−1.02) | (−2.74) | (−.45) | | (−1.64) | (.08) | (−1.60) | | |
| 3‡ | .79 | .23 | .23 | .26 | ... | ... | ... | ... | .26 | .22 | .15 | ... | 330 |
| | (10.73) | (17.00) | (6.75) | (5.87) | | | | | (6.04) | (5.80) | (3.09) | | |

NOTE.—Assoc., Recip., Sub., Nonsub. are abbreviations for association, reciprocal, subsidiary and nonsubsidiary, respectively. Group is a dummy variable indicating group membership. $F$ indicates the value of the $F$-statistic under the null hypothesis that all coefficients are identically equal to zero. $N$ represents the sample size. Classifications Assocation Owned, Mutual Owned, Single Owner (Sub.)/(Nonsub.), Closely Held (Sub.)/(Nonsub.), and Widely Held (Sub.)/(Nonsub.) are all stock companies. Parentheses contain $t$-statistics.

* These regressions omit the intercept term to avoid singularity in the regression matrix. Alternatively, we could have omitted one of the ownership classification dummy variables. The coefficients reported in regressions 1 and 3 are the mean ratios for the respective ownership classification and sample.

† This row and the one below contain coefficient estimates and $t$-statistics (respectively) for interaction effects between group membership and the indicated ownership classification.

‡ This regression excludes subsidiary companies and group members.



Regression 3 employs the subsample of 330 companies that excludes subsidiaries and group members.

Each regression indicates that reinsurance is a significant fraction of total business for all ownership classes. The extremes appear to be the Lloyd's associations in regression 1, with a mean ratio of .88, and the widely held (nonsubsidiary/nongroup) stock companies from regression 3, where the mean ratio is .15. From regression 2, group membership increases the reinsurance ratio by 27% for widely held stock subsidiaries. Three (out of 10) interaction terms between group membership and ownership classification have statistically significant negative coefficients, but in all three the interaction term coefficient point estimates are less (in absolute value) than the significant positive coefficient estimate on group membership. Thus, group membership is generally associated with larger apparent usage of reinsurance. Regression 2 also suggests, at least weakly, that subsidiaries purchase more reinsurance than do nonsubsidiaries, even controlling for group membership. The nonsubsidiary ownership classifications all indicate lower mean ratios than their respective subsidiary counterparts. An additional implication from comparing mean ratios from regression 3 with those from regression 1 is that subsample mean ratios appear close to those from the full sample (except for the Lloyd's associations). Thus, at least based on the mean, the subsample appears representative of the full sample for the nonsubsidiary ownership classes; group membership appears most important for subsidiary companies.

Table 5 contains pairwise comparisons of mean and median reinsurance activity for the alternate ownership classes. The tests are two-sample $t$-tests with unequal variances for the means and Wilcoxon two-sample (Mann-Whitney) tests for the medians (see Siegel and Castellan 1988). Panel A of table 5 reports tests for the full sample of 1,276 companies, and panel B reports the results for the subsample of 330 nongroup/nonsubsidiary companies. While the mean and median tests provide reinforcing implications, the significance level typically is lower for the median.

The table 5 evidence suggests strongly that Lloyd's associations spend a larger proportion of their premiums on reinsurance than any other ownership class. Also, there is weak evidence that widely held (nonsubsidiary) stock companies spend a smaller proportion of their premiums on reinsurance than other ownership classes. These results are consistent with our ownership-structure hypotheses. Finally, from panel A, there is evidence that subsidiaries purchase more reinsurance than nonsubsidiary firms. For example, except for the Lloyd's, mutual-owned stock companies have larger mean and median reinsurance ratios than any other nonsubsidiary ownership class. This increase in the reinsurance ratio for subsidiaries and group members is consistent with the use of group business pooling arrangements. If our data al-

**TABLE 5** Pairwise Comparisons of Mean/(Median) Reinsurance Activity (Ratio of Reinsurance Premiums Ceded to Total Business Premiums) for Alternative Ownership Classifications (Two-Sample *t*/[Wilcoxon Two-Sample Z])

A. Total Sample of 1,276 Property/Casualty Insurance Companies

| Ownership Classifications | Mutual | Reciprocal Association | Association Owned | Mutual Owned (Subsidiary) | Single Owner (Subsidiary) | Closely Held (Subsidiary) | Widely Held (Subsidiary) | Single Owner (Non-subsidiary) | Closely Held (Non-subsidiary) | Widely Held (Non-subsidiary) |
|---|---|---|---|---|---|---|---|---|---|---|
| Lloyd's association | 15.75 (9.28) | 11.94 (7.27) | 9.89 (6.18) | 9.16 (6.83) | 10.80 (7.56) | 5.28 (4.57) | 10.13 (7.46) | 8.85 (5.80) | 11.39 (6.58) | 13.75 (6.43) |
| Mutual | | −.44 (.50) | −.79 (−.14) | −5.46 (−4.30) | −2.38 (−1.29) | −4.72 (−4.71) | −11.08 (−8.08) | .03 (1.11) | −.66 (−.59) | 1.98 (1.90) |
| Reciprocal association | | | −.38 (−.34) | −3.55 (−3.04) | −1.42 (−1.08) | −3.87 (−3.77) | −5.22 (−4.20) | .26 (.63) | −.19 (−.86) | 1.80 (1.16) |
| Association owned | | | | −2.56 (2.11) | −.81 (−.57) | −3.21 (−3.10) | −3.59 (−2.93) | .52 (.96) | .21 (−.31) | 1.89 (1.39) |
| Mutual owned (subsidiary) | | | | | 2.13 (2.03) | −1.42 (−1.30) | −.94 (−.82) | 2.68 (2.61) | 3.22 (2.16) | 5.49 (3.39) |
| Single owner (subsidiary) | | | | | | −2.87 (−2.72) | −3.53 (−3.24) | 1.24 (1.40) | 1.17 (.45) | 3.26 (2.04) |
| Closely held (subsidiary) | | | | | | | 1.01 (1.06) | 3.31 (2.89) | 3.65 (3.09) | 5.21 (4.07) |
| Widely held (subsidiary) | | | | | | | | 3.44 (3.11) | 4.69 (2.85) | 7.66 (4.21) |
| Single owner (nonsubsidiary) | | | | | | | | | −.39 (−.98) | .99 (.04) |
| Closely held (nonsubsidiary) | | | | | | | | | | 1.92 (1.77) |

B. Sample of 330 Nongroup/Nonsubsidiary Property/Casualty Insurance Companies

| Ownership Classifications | Mutual | Reciprocal Association | Association Owned | Single Owner (Nonsubsidiary) | Closely Held (Nonsubsidiary) | Widely Held (Nonsubsidiary) |
|---|---|---|---|---|---|---|
| Lloyd's associations | 7.12 (4.25) | 6.61 (3.70) | 5.39 (3.29) | 5.34 (3.19) | 6.63 (3.81) | 7.55 (3.68) |
| Mutual | | .06 (.30) | −.40 (.42) | .41 (.74) | .24 (.49) | 2.26 (2.02) |
| Reciprocal Associations | | | −.38 (.19) | −.39 (.26) | .14 (.01) | 1.68 (1.38) |
| Association Owned | | | | −.01 (.20) | .48 (.03) | 1.55 (1.02) |
| Single owner (nonsubsidiary) | | | | | .49 (−.20) | 1.54 (.46) |
| Closely held (nonsubsidiary) | | | | | | 1.49 (1.20) |

NOTE.—The two-sample *t*-test assumes unequal variances. The Wilcoxon two-sample test is a median test. This test is also referred to as the Mann-Whitney test (see, e.g., Siegel and Castellan 1988).

lowed us to distinguish between reinsurance purchases where the rein-
surer is another member of the same group and purchases where the
firm obtains reinsurance from outside the group, then we could control
more effectively for these group problems. However, without that
data, our most effective tests must focus on the subset of 330 non-
group, nonsubsidiary firms.

*Size concentration and Best's rating.* In table 5, we examine the
impact of ownership structure on reinsurance purchases without con-
trolling for other factors. In table 6 we examine the effects of size, line-
of-business concentration, geographic concentration, default risk, and
organization/ownership structure on reinsurance purchases. We report
three regressions: regression 1 contains the size, line-of-business con-
centration, Best's rating, and geographic concentration measures as
independent variables; regression 2 adds the ownership class dummy
variables; and regression 3 adds the percentage of business written in
each of 23 lines of insurance as control variables.[20]

We omit the widely held stock ownership class dummy variable from
regressions 2 and 3 to avoid singularity. Thus, the $t$-statistics for own-
ership class variables test whether the mean ratio for the ownership
class is different from the widely held stock mean ratio. We also exam-
ine pairwise tests of the equality of estimated ownership class coeffi-
cients from regression 3 and report these $F$-statistics at the bottom of
table 6.

Again, Lloyd's appear to have a larger ratio of reinsurance to total
business premiums than other ownership classes: the $F$-statistics from
table 6 range from 17.08 to 30.96 and the $t$-statistic of 5.80 from regres-
sion 3 are all highly significant. However, part of this difference may
reflect the fact that Best's data do not differentiate well between insur-
ance business where the cash flows are more uncertain and where they
are not.

There is also strong evidence that widely held stocks reinsure less
than closely held, single-owner, and association-owned stocks: the $t$-
statistics for these comparisons are significant at least at the .05 level.
Moreover, there is weak evidence that the ratio of reinsurance to total
business is smaller for widely held stocks than for any other ownership
class: the widely held/reciprocal test has the largest $p$-value of any
widely held stock test, and it is .147. There is also weak evidence that
single-owner stocks reinsure more than mutuals and reciprocals.

20. In table 6, regressions 2 and 3 imply restrictions on the interactions among the independent variables. As a specification check, we employ an analysis of covariance where our ownership classification variables are interpreted as treatments and other variables are covariates. Tests for interaction effects between the ownership classes and line-of-business concentration, size, geographic concentration, and Best's rating indicate they are not significant at the .10 level.



TABLE 6 Cross-sectional Regressions of Reinsurance Activity (Ratio of Reinsurance Premiums Ceded to Total Business Premiums) on Ownership Classification Dummy Variables, Size, Line-of-Business Concentration (HERF), Geographic Concentration (LICENSE), and Best's Rating for Nonsubsidiary, Nongroup Subsample of 330 Property/Casualty Insurance Companies

A. Variables

| Regression No. | Intercept | Lloyd's Assoc. | Mutual | Recip. Assoc. | Assoc. Owned | Single Owner | Closely Held | SIZE | Business Concentration (HERF) | Geographic Concentration (LICENSE) | Best's Rating | $F$ | $\bar{R}_2$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1.211 | | | | | | | −.055 | −.079 | −.003 | −.015 | 13.80 | .13 |
| | (8.84) | | | | | | | (−6.60) | (−1.75) | (−3.00) | (−2.76) | | |
| 2 | .967 | .547 | .115 | .094 | .114 | .058 | .031 | −.049 | −.013 | −.003 | −.013 | 11.10 | .23 |
| | (6.57) | (6.31) | (2.36) | (1.66) | (1.87) | (.94) | (.53) | (−5.91) | (−.29) | (−3.21) | (−2.38) | | |
| 3* | 1.053 | .475 | .084 | .077 | .112 | .149 | .131 | −.033 | −.129 | −.002 | −.014 | 7.26 | .39 |
| | (5.58) | (5.80) | (1.81) | (1.45) | (1.96) | (2.44) | (2.22) | (−3.62) | (−2.43) | (−2.16) | (−2.41) | | |

B. Tests of the Equality of Ownership Classification Coefficients for Regression 3 ($F$-Statistic/$P$-Value)

| Ownership Classification | Mutual | Recip. Assoc. | Assoc. Owned | Single Owner | Closely Held |
|---|---|---|---|---|---|
| Lloyds assoc. | 30.96 | 28.25 | 20.79 | 17.08 | 19.25 |
| | (.0001) | (.0001) | (.0001) | (.0001) | (.0001) |
| Mutual | | .04 | .45 | 1.98 | 1.10 |
| | | (.8516) | (.5012) | (.1609) | (.2951) |
| Recip. assoc. | | | .52 | 1.91 | 1.18 |
| | | | (.4732) | (.1680) | (.2774) |
| Assoc. owned | | | | .40 | .11 |
| | | | | (.5288) | (.7424) |
| Single owner | | | | | .12 |
| | | | | | (.7242) |

NOTE.—Assoc. and Recip. are abbreviations for association and reciprocal. Classifications association-owned, single owner, closely held, and widely held are stock companies. The SIZE variable is the log of total admitted assets. HERF is the Herfindahl line-of-business concentration measure, and LICENSE is the negative of the number of states in which the businesses are licensed. $F$ indicates the value of the $F$-statistic under the null hypothesis that all coefficients are identically equal to zero $\bar{R}_2$ is the adjusted $R^2$.

* Regression 3 controls for the percentage of business written in individual lines of insurance. Thus, this regression includes an additional 23 variables, one for each of the lines of insurance in table 3. We omit the coefficient estimates and $t$-statistics for those variables; however, they are available from the authors.

The evidence in table 6 also suggests that size, line-of-business concentration, geographic concentration, and Best's rating have a significant negative impact on the demand for reinsurance. The significant coefficient on our geographic concentration variable implies that the real-service efficiency argument (which implies a negative coefficient) is quantitatively more important than the sum of the other effects through taxes, expected bankruptcy costs, and investment incentives (which all imply positive coefficients). Real-service efficiencies also imply negative coefficients for size and line-of-business concentration, but for these variables the signs of the other effects (except for bankruptcy costs on size which is also negative) are ambiguous. Thus, our tests do not allow us to identify a dominant factor in explaining the signs and significance of size and line-of-business concentration as explanatory variables for reinsurance demand.

We do not report coefficients and $t$-statistics in table 6 for the 23 line-of-insurance variables that are included in regression 3 since none of our hypotheses are related to the magnitudes of any of the coefficients. However, the importance of including these control variables should be obvious from an examination of the reported statistics; for example, the coefficient on line-of-business concentration is insignificant unless variables controlling for the percentage of business written in individual lines are included. More directly, the importance of including these variables can be assessed by observing that the adjusted $R^2$ in regression 3 is substantially larger than in regression 2 (.39 vs. .23) where the variables are not included.

## V. Conclusions

The examination of reinsurance purchases by 1,276 property/casualty insurance companies provides evidence that ownership structure matters. Generally, the less diversified the owners' portfolios, the greater the reinsurance purchases. Thus Lloyd's reinsure most, while widely held stocks reinsure least. Moreover, subsidiary and group relations affect the demand for reinsurance. Subsidiaries and group members reinsure more (although our data do not allow us to distinguish between intragroup transactions and reinsurance transactions with external companies).

We also provide evidence that size, credit standing, and geographic concentration reduce the demand for reinsurance and weak evidence that line-of-business concentration reduces reinsurance demand, as well. Our estimated negative effect of geographic concentration suggests that the real-services argument is quantitatively important. The substantial explanatory improvement we obtain in our cross-sectional regression, by including the percentages of business written in the individual lines, indicates the importance of controlling for variation in



operating characteristics across lines of business for explaining variation in reinsurance purchases.

There are some potentially important limitations to our analysis: (1) The power of our tests is reduced by our lack of information about the tax status of individual firms. (2) More powerful tests could be designed if we had independent estimates of the operating cash flow volatilities of each line. (3) Our data are aggregated into lines of insurance; while within lines, policies are undoubtedly heterogeneous in their riskiness. With more detailed information about the risks of the specific policies sold by particular firms, tests with greater power are possible. (4) Our data do not distinguish between reinsurance purchases where the reinsurer is another member of the same group and purchases where a group member obtains reinsurance from an external reinsurer. With such data, more powerful tests employing our larger sample of firms are possible.

Our original motivation for this article was to provide evidence on the corporate demand for insurance. For industrial firms, the signs of the partial effects of ownership structure, credit standing, and size should be direct corollaries of our insurance industry results. Similarly, we expect the industrial firms' insurance demand to vary with the operating characteristics of their business. However, we doubt that the demand for insurance by noninsurance firms is related to their geographic or line-of-business concentration. If data on industrial firms' insurance purchases were available, then other proxies for real-service efficiencies, tax benefits, bankruptcy costs, and investment incentive effects would be required.

## References


Arrow, K. J. 1964. *Essays in the Theory of Risk-Bearing*. Amsterdam: North-Holland.

Best, A. M. 1982. *Best's Insurance Reports—Property Casualty*. Oldwick, N.J.: A. M. Best Company.

Bickelhaupt, David L. 1983. *General Insurance*. Homewood, Ill.: Richard Irwin.

Hendershott, Patric, and Koch, Timothy W. 1980. The demand for tax-exempt securities by financial institutions. *Journal of Finance* 35 (June): 717–27.

Huberman, Gur; Mayers, David; and Smith, Clifford. 1983. Optimal insurance policy indemnity schedules. *Bell Journal of Economics* 4 (Autumn): 425–26.

Johnson, Herb, and Stultz, René. 1987. The pricing of options with default risk. *Journal of Finance* 42, no. 2 (June): 267–80.

Main, B. G. M. 1983. Corporate insurance purchases and taxes. *Journal of Risk and Insurance* (June), pp. 197–223.

Mayers, David, and Smith, Clifford. 1981. Contractual provisions, organizational structure, and conflict control in insurance markets. *Journal of Business* 54 (July): 407–34.

Mayers, David, and Smith, Clifford. 1982. On the corporate demand for insurance. *Journal of Business* 55, no. 2 (April): 281–96.

Mayers, David, and Smith, Clifford. 1987. Corporate insurance and the underinvestment problem. *Journal of Risk and Insurance* (March), pp. 45–54.

Mayers, David, and Smith, Clifford. 1989. Executive compensation in the life insurance industry. Working paper. Rochester, N.Y.: University of Rochester.

# CHAPTER 4

## THE PRICING OF INSURANCE CONTRACTS

An insurance contract is a promise to pay future losses that may arise. What is the value of that promise? In one sense, the value is indicated by the price people are willing to pay for that promise. In another sense, the value is indicated by the price at which others are willing to sell the promise. Price, it seems, provides a measure of what the insurance policy is worth. In this chapter, we will start to examine insurance prices or premiums. Amongst the questions that will be raised are the following:

How will the insurance company set its prices?

How will these prices vary between policyholders with differing loss expectancies?

In setting the price, how do we cope with the fact that losses are uncertain and will arise sometime in the future?

What price would prevail if the insurance market is highly competitive?

What is the minimum price necessary to keep an insurance company in business?

What effect will the investment income of the insurer have on the prices of its insurance policies?

We will address these questions in the following sequence. First, we present the general principles that permit the firm to set its prices to maximise profits. The profit maximising prices will then be shown to discriminate between individual policyholders according to their respective loss expectancies. Such discrimination is in part supported by, and in part constrained by, statute law. The general pricing problems for mutual and reciprocal firms will then be addressed. The chapter continues by looking at the practical systems adopted by insurers (classification and individual rating) to derive appropriate premiums.

Premium or price setting for the insurer gives rise to another problem. The various cash flows in a contract are uncertain and arise at different points in time. Discounted cash flow techniques are presented to value such cash flows. These techniques are then adapted to insurance pricing.

To set the wheels in motion, we will consider a very simple problem of how any firm (not necessarily an insurance firm) would seek to set its prices.

### PRICING FOR THE PROFIT MAXIMISING FIRM

#### (a). An Illustration for the Non-Insurance Firm.

Our firm produces widgets.[1]. The firm, we assume, is in business to maximise its profit which is defined to be the difference between revenue and cost. The decisions we shall address concern the quantity of widgets to produce and the price at which to sell. In fact, these are not separate decisions at all. If the price is set by the firm, the customers will decide how much they wish to purchase at that price. In other words, the firms faces a demand curve for its product. The more expensive are widgets, the less people will be inclined to purchase. But the firm can increase the volume of its sales by lowering the price. The demand curve is shown in figure 1.

Table 1 shows the hypothetical information for the firm. First, notice that the demand curve slopes downwards as in figure 1. The firm can raise its price but it will suffer a decline in the quantity it can sell. The next feature of interest is the cost of production. Notice that the total cost of production rises, quite naturally, as the number of units produced rises as shown in column 2. But the rise in costs is not proportional to the number of units produced. You may wish to calculate the average cost per unit by dividing the total cost by the number of units. This will reveal that average costs declines for a while as output increases then starts to rise again when output exceeds six units.[2] Now the marginal (or incremental) cost shows the amount by which costs increase as we add another unit to output. Now let us look at the revenues and profit. The total revenue is simply the total income from sales; i.e. price times quantity sold. The marginal (incremental) is simply the addition to total revenue from adding another unit to output.

The firm, by assumption wishes to maximise profit. Of course, we can see the answer straight away in the final column. If six units are sold at a price of 9.3 per unit, the profit is at a maximum at 15.8. But let us examine the pattern in marginal cost and marginal benefit that leads to this profit maximising output level. Look, for example, at the quantity four units and consider what would happen to marginal cost and marginal revenue if we increased the quantity from four units to five units. Marginal cost is $5 and marginal revenue is $8. More has been added to revenue than to cost therefore the profit increases by $3 as shown in the final colummn. Clearly, there is benefit from increasing output from 4 to 5 units. If we increase output from 5 to 6 units, again marginal revenue exceeds marginal cost, by $0.8, and this increase in output increases profit. But increasingoutput again from six to 7 units adds $7 to costs but only $5.8 to revenue; profit consequently falls by $1.2. The pattern is clear. We should continue to increase output as long as marginal revenue exceeds marginal cost but only up to the point at which they are equal. <u>The profit maximising output occurs where marginal cost equals marginal revenue.</u> Increases in output beyond this point reduce profit since marginal costs exceed marginal revenue. This is a well known decision rule in economics.

----------

1. This is the technical name for a device that prevents economics textbooks from falling apart

2. This pattern shows increasing returns to scale up to six units and decreasing returns to scale thereafter.

These ideas are illustrated in Figure 2. In the top part of the diagram, the marginal cost and marginal revenue schedules are graphed from the data in table 1. Also shown is the price or demand line. The intersection of the marginal costs and revenue lines does not occur exactly at 6 units (there is a small issue of discrete units of production). The lower part of the Figure plots the profit at different levels of output. This peaks at 6 units which is directly adjacent to the intersection of the marginal cost and revenue curves. We will now use these ideas to look at insurance and in doing so we will repeat the decision rule. Profit is maximised at the level of output at which marginal cost equals marginal revenue.

(b). Price Discrimination for the Insurance Firm.

We happened to be talking about widgets in the above example. The analysis would be the same if we had been discussing the pricing of automobiles, radios, fertiliser, bank deposits or insurance policies. As long as we measure price and output correctly, the golden rule (marginal cost = marginal revenue results in a profit maximising solution) applies in each case.

Figure 1

The Demand Curve for Widgets

```
price
 |
 |  \
 |   \
 |    \
 |     \
 |      \
 |       \demand
 |        \
 -------------->quantity
```

Table 1

Calculation of the Profit Maximising Price and Output

| Units Sold | Total Cost $ | Incremental or Marginal Cost $ | Price $ | Total Revenue $ | Incremental or Marginal Revenue $ | Profit $ |
|---|---|---|---|---|---|---|
| 1 | 10 | 10 | 12 | 12 | 12 | 2 |
| 2 | 18 | 8 | 11.3 | 22.6 | 10.6 | 4.6 |
| 3 | 24 | 6 | 10.7 | 32.1 | 9.5 | 8.1 |
| 4 | 29 | 5 | 10.25 | 41 | 8.9 | 12 |
| 5 | 34 | 5 | 9.8 | 49 | 8 | 15 |
| 6 | 40 | 6 | 9.3 | 55.8 | 6.8 | 15.8 |
| 7 | 47 | 7 | 8.8 | 61.6 | 5.8 | 14.6 |
| 8 | 55 | 8 | 8.4 | 67.2 | 5.6 | 12.2 |
| 9 | 64 | 9 | 8 | 72 | 4.8 | 8 |
| 10 | 74 | 10 | 7.6 | 76 | 4 | 2 |

Now let us use these ideas in the following problem. There are two groups of potential clients for the insurance firm; high risks and low risk. We may think of the high risks as bad drivers seeking automobile insurance and who have a high probability of generating a loss. Low risks, by contrast, are good drivers who have a low loss expectancy. The costs of doing business for

Figure 3

Profit Maximisation & Price Discrimination

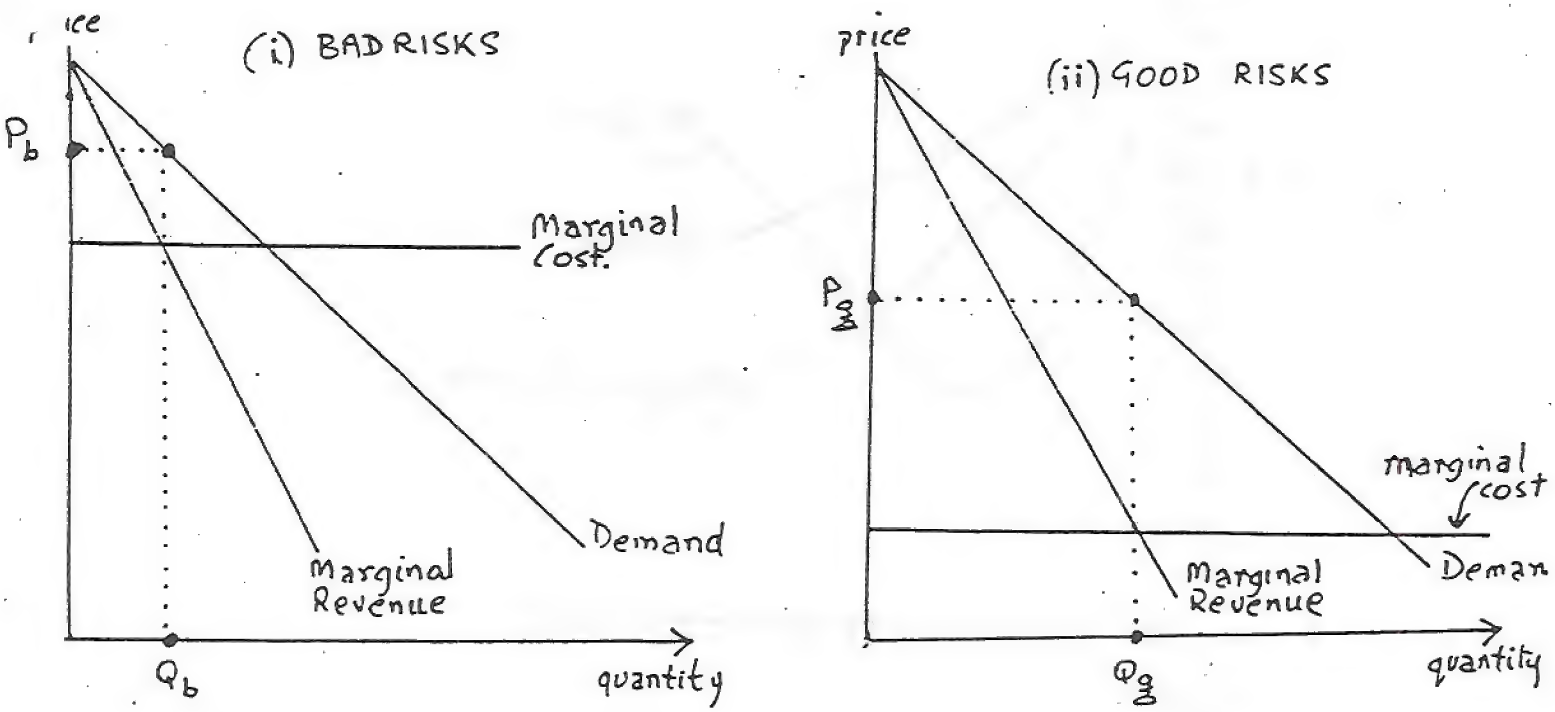
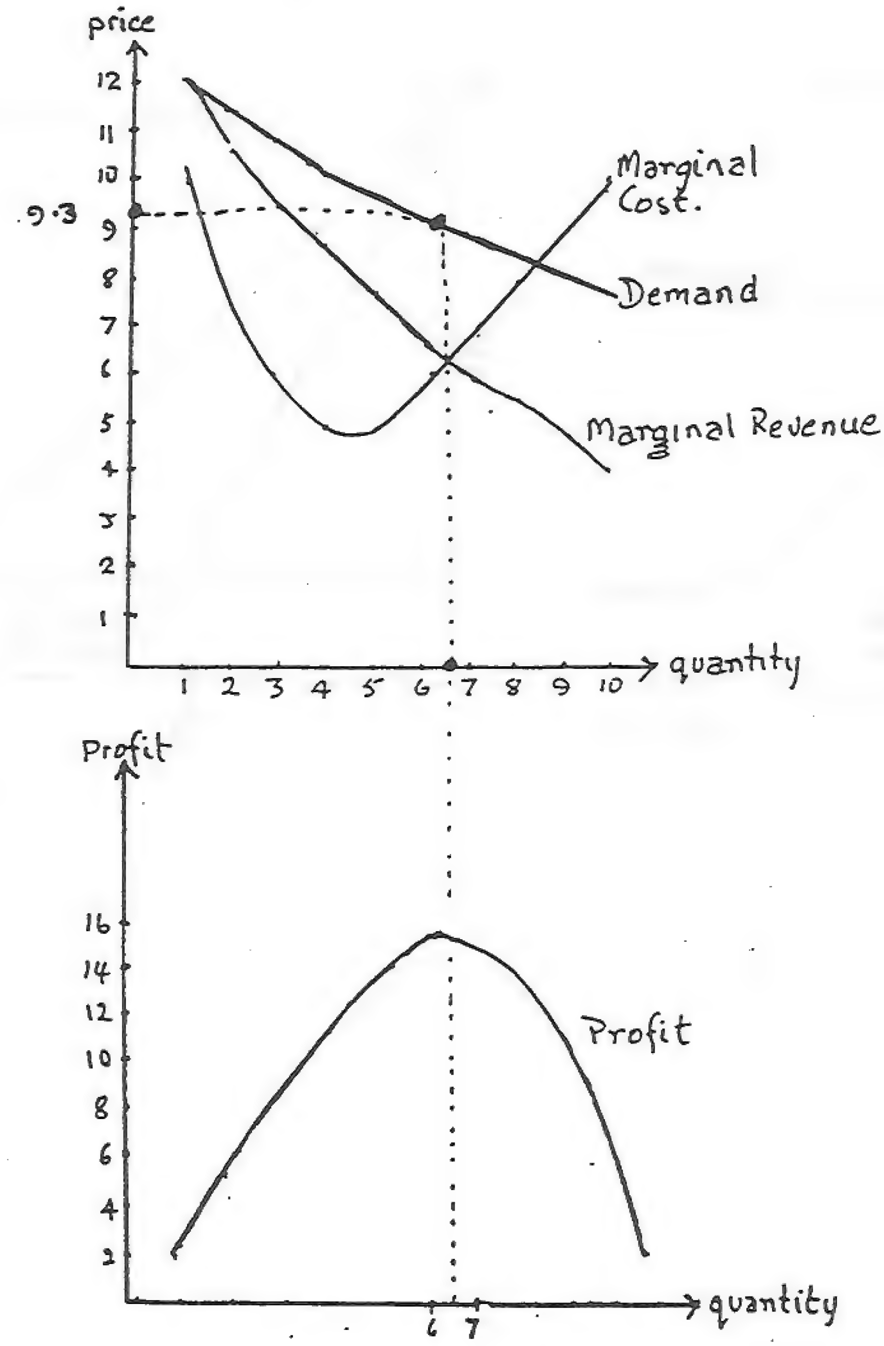


Figure 2

Illustration of Marginal Cost = Marginal Revenue Rule for Profit Maximisation



the insurance firm mainly comprise the costs of settling claims. Thus the bad risks will impose higher costs on the insurance firm. This difference is represented in Figure 3. The left side shows the bad risks and the right side shows the good risks. Marginal cost curves are shown revealing the bad risks to impose considerably higher costs on the insurer than the good risks. The demand curves for each group and the marginal revenue curves also are shown. For simplicity, these curves are identical; at a given ratio of premiums to expected losses each group would buy an equivalent amount of insurance.

Given these two groups of potential clients, what is the profit maximising strategy for the insurer. The marginal cost = marginal revenue rule can be applied separately to each market. This shows that quantities $Q_b$ and $Q_g$ should be sold to the bad and good risks respectively. But at what prices? The demand curves show that $Q_b$ should sell at price $P_b$ and that $Q_g$ should sell at $P_g$. Separate prices should be charged to the two groups in order to maximise profit. The bad risks should be charged a higher price than good risks. To get a different perspective on this consider the effect of charging the same price for the two groups. Intuition suggests that a common price for the two groups would lie between $P_b$ and $P_g$, say at P*. But we know that any price other than $P_b$ would yield less than the maximum possible profit in the bad risk market and that any price other than $P_g$ would yield less than the maximum profit in the good risk market. Conseqüently, P* yields lower profits in both markets than the separate prices.

(c). Is Price Discrimination Supported or Restrained in Law?

This analysis shows that insurers have a profit incentive to charge their clients different prices according to differences in their respective loss expectancies. This principle may be called actuarial discrimination. It is not the same thing as discrimination in the civil rights sense though there may be some overlaps. Say, for example, men have a higher expected value of loss than women. This is generally the case when considering mortality rates for life insurance or in considering the accident probabilities of young people seeking automobile insurance. The actuarial considerations would imply that insurers would seek higher premiums for men than for women. But a civil rights law might mandate the same premium. There is clearly a conflict. But the conflict is not simply between the profit of the insurer and social equity. There are other considerations. A common premium would create a subsidy from low risk women to high risk men. Moreover, this subsidy will make insurance relatively expensive for the low risk group and cheap for the high risk group (relative, in this case, to the expected value of loss for each group). Thus an adverse selection problem arises that results in a strong demand for insurance in the high risk group and a weak demand in the low risk group. In the extreme, this process might drive the good risks out of the market.

Societies express social values in their laws. With respect to insurance pricing there are conflicting values. Most States generally support the concept of actuarial discrimination. For example, consider the following,

from Pennsylvania[3]

> Risks may be grouped by classification for the establishment of rates and minimum premiums. Classification rates may be modified to produce rates for individual risks in accordance with rating plans which establish standards for measuring variations in hazards or expense provisions, or both. Such standards may measure any differences among risks that can be demonstrated to have a probable effect on losses or expenses.

The legislative endorsement of actuarial discrimination in Pennsylvania is not unusual. Most states have similar law. This is explicable simply by the view that bad risks should "pay their way" and not be supported by subsidies from the good risks. This objective is often referred to as the requirement that insurance rates be "not unfairly discriminatory". This phrase is widely used to convey the notion that

> people with the similar expected losses and imposing similar expenses on the insurance firm should be charged similar prices
>
> people with different expected losses and imposing different expenses on the insurance firm should be charged different premiums reflecting these differences

The term "not unfairly discriminatory" conveys the same idea as "actuarial discrimination". The following is an extract from Wisconsin Law[4]

> One rate is unfairly discriminatory in relation to another in the same class if it clearly fails to reflect equitably the differences in expected losses and expenses. Rates are not unfairly discriminatory because diffent premiums result for policyholders with like loss exposures but different expense factors, or like expense factors but different loss exposures, so long as the rates reflect the differences with reasonable accuracy.

But overlying this view of discrimination is another which not only affects insurance contracts, but extends across all economic activity. The use of factors such as race, religous preference, sex, age or marital status to distinguish people when making economic contracts has been thrown into question. Most agreement is found in the use of race and religous preference where Federal and State statutes prohibiting discrimination abound. The other factors, sex, age and marital status are all sensitive areas in which one finds selective legislaion but widespread debate. At this time these factors are widely used in insurance ratemaking but are under challenge.

----------

3. Pa. Stat. Ann. tit40, Sec.1183(c)(1971). A good presentation of these issues is given in McNamara [1985].

4. Wis. Stat. Ann. Sec 625.11(4)(1980)

(d). Price Discrimination, Moral Hazard and Adverse Selection

In the previous chapter, two potential problems for insurance markets were discussed; moral hazard and adverse selection. In extreme cases, each of the problems could prove fatal to the insurance market. But in less than extreme cases moral hazard represents a resource cost that is borne by all policyholders and adverse selection makes insurance either less available, or more expensive, for good risks. Are there any antidotes for these problems. In principle, there are. If policies are each priced according to loss expectancy, then both problems should disappear.

Moral hazard arises when the presence of the insurance policy deters the policyholder from those activities that would reduce the probability or severity of loss. There is little incentive for safety if someone else foots the bill for the loss. But premium discrimination implies that loss reduction and safety is rewarded in the premium. The policyholder does have an incentive to instal a sprinkler or to drive safely if such behavior will reduce premiums either now or in the future. Consider a simple example;

**********************************ILLUSTRATION**********************************

Scene 1. The Board Room

The Pyro Nuts Co is deciding to instal a sprinkler that costs $250,000. The expected lifetime of the sprinkler system is estimated to be ten years. This will reduce the expected value of future losses to the firm from an annual estimated cost of $250,000 per year to an estimated $150,000. On the face of it, this seems a good investment. Spending $250,000 now and making an annual saving of $100,000 for the next ten years. But the firm is full insured and the directors reason that it is unwise to spend $250,000 so that the insurance company's claim costs can be reduced.

Scene 2. Still in the Board Room

In the nick of time, the risk manager who had been on the factory floor fighting dangerous housekeeping practices, lax safety procedures and oiling the automatic fireproof doors, bursts into the board room holding the firm's fire insurance policy high in the air. "Did you know the Prudent will reduce our premiums by $100,000 if we instal sprinklers?"

Scene 4. Coming out of the Board Room after having decided to instal the Sprinkler System

The Chairman puts his arm around the risk manager's shoulder and says' "You know Jenkins, I've had my eye on you for some time. I would like you to come around for dinner at the weekend and meet the family."

********************************************************************************

This example is a little outrageous but it does illustrate that insurance premiums which are related to loss expectancy will convey incentives to the policyholder for loss prevention. But actuarial discrimination also reduces the adverse selection problem. Adverse selection arises when a common premium is charged to policyholders having different loss expectancies. Under these circumstances, insurance will be relatively unattractive to the low risk individual but relatively cheap to the high risk individual. For the latter, the premium is subsidised. This practice will induce many low risk

individuals not to purchase insurance leaving the insurer with a disproportionate number of bad risks. But if the insurer is able to, and is permitted to, distinguish between the individuals and charge each a separate premium, the problem disappears.

The effectiveness of insurance pricing systems in combatting moral hazard and adverse selection depends on the ability of the insurer to measure differences in loss expectancies between different policyholders. Presently, we will address the practical issues in rating. In the meantime, the pricing problem for a mutual is considered.

## PRICING FOR A MUTUAL INSURANCE FIRM

The investors in a stock insurance company invest in order to secure the most attractive rate of return on their capital investment. Profit maximisation is generally consistent with this objective. But what about the mutual or the reciprocal insurance firm? There are no shareholders, the equity being supplied by the policyholders. If these firms make profits, they are simply taking transfering money from the policyholders back to the policyholders; they pay extra in premiums only to receive it back in the form of dividends.

Since the existing policyholders own the firm, then a plausible objective for the mutual is that it maximise the welfare of the existing policyholders. These policyholders are interested in having adequate insurance protection at an attractive price. But since the mutual may pay a dividend, the effective price to the policyholder should net out any expected dividend. Unfortunately, we are not aware of any formal research on the pricing strategies of mutuals, but we should be able to make some rather general statements. Consider the following thoughts.

all equity to the mutual must be provided by the policyholders

the policyholders do not want the firm to become insolvent since any outstanding claims would not be fully paid

the higher the level of equity, the lower than probability of insolvency

the policyholder does not wish to pay a large premium at the beginning of the contract

nevetheless, if premiums as a whole do prove to be more than adequate to pay losses and to provide for future equity needs, a dividend will be paid to the policyholders

low risk policyholders will not wish to subsidise high risk policyholders

These thoughts seem to lead to the following conclusions. Mutuals will price their policies to cover the expected loss but the price must also include a contribution to the equity of the firm. The firm will pay dividends when it considers the equity to be more than sufficient to provide a

reasonable probability that the firm will remain solvent. Policyholders will be willing to pay a reasonable contribution to equity to keep the firm financially sound. But if the firm charges too much up front, it will lose its policyholders who will probably now find it preferable to go to stock insurance company. The we would expect the mutual to charge premiums that were sufficient to cover expected claims and provide a reasonable margin for equity. The premium must perform a "balancing act" between the need to keep the price competitive and the need to provide sufficient equity. The mutual would also be expected to provide premiums that discriminated between high and low risks subject, of course, to the constaints of civil rights legislation.

## RATING SCHEMES

Considerations of both economics and fairness indicate that insurance prices should discriminate according to differences in loss and expense factors within the constraints imposed by civil rights legislation. Insurers use two main devices to allocate different premiums to different policyholders; these are classification and individual experience.

### (a) Classification and Insurance Premiums

For many purposes in life we stress the similarities between people and for other purposes we stress differences. When putting together an electoral register, we will count all adults who are citizens (any differences are irrelevant for this purpose). When imposing income taxes, we require that all people whose defined income (net of allowable deductions) is in a certain range, to pay similar taxes. But for other purposes, the differences are vital. It is not sufficient to know that a white male in his early 30's committed the crime, we wish to know which particular white male. Or, if I am "choosing" a wife, it is unecessarilly democratic to classify all single females in the required age group as equally suitable.

In setting insurance premiums, we would ideally like to know the expected costs which each individual imposes on the insurance firm, both expenses and claims. But this is not a practical proposition. An extensive examination cannot be made of every individual. This would be prohibitively expensive. A compromise usually is reached by classifying policyholders into sub groups of fairly similar individuals. This is achieved by identifying observable features of the policyholders that are thought to be related to their loss expectancies. For example, mortality rates are known to be strongly related to age and sex. Thus life insurance premiums are set by classifying policyholders into subgroups comprising the same age and sex. The premium is then set for the subgroup usually based on the aggregate loss experience for that subgroup. This is illustrated below in the following extract from a life insurance classification scheme.

| \ age<br>sex \ | ............. | 21 | 22 | 23 | 24 | 25 | 26 | ......> |
|---|---|---|---|---|---|---|---|---|
| male | ............. | | | | x | | | ......> |
| female | ............. | | | | | | | ......> |

Each cell in the matrix represents a subgroup. For example the cell marked "x" comprises 24 year old males. Within this class, individuals are treated similarly.

Similar classification schemes are used in other lines of insurance. For automobile insurance, the insurer may use the following set of variables (actual practice does vary)

age, sex, marital status, garage location, use of vehicle and mileage.

and, for fire insurance on commercial property;

use of property (industrial classification), structure, fire protection (sprinklers, fire alarms etc), location, materials stored etc..

Classifying risks in this way, represents an attempt to provide a rating system that is roughly discriminatory across individual ppolicyholders and yet which is economic to operate and makes best use of statistical data in inferring loss probabilities. To be useful, the factors used for classification must be correlated with the incidence of loss and must be observable. Take age and sex in life insurance. Both age and sex are easily observable (fortunately, in the interests of privacy, we do not need direct observation on either variable). Most of us would accept that age itself is causally related with mortality rates. Sex may not have any direct causal relationship with mortality. But the association between sex and mortality rates is very strong. Perhaps sex is acting as a proxy for other factors in the causal relationship such as differences in the lifestyles of men and women. But these underlying factors, whatever they may be, may be more difficult to observe. Thus sex serves as a useful practical rating factor. In general, it is not necessary to establish a causal relationship to make a rating factor useful.

(b). Individual rating

Sometimes, individual features of a risk that are not necessarily associated with other groups of policyholders, are used in insurance pricing. A good example is experience rating. Having established a premium for an insured by use of some system of classification, the premium may be adjusted in relation to the individual policy claim experience. Two branches of insurance in which experience rating is extensively used are automobile insurance and workers' compensation insurance. For example, in automobile insurance, many insurers have a set of internal rules to increase the premium if the individual policyholder has a certain number of claims within a given period. Alternatively, the premium may be reduced if the loss experience is favorable.

In Europe, these "no claims discount" schemes are even more extensively used and the premium adjustment rules are made known to the policyholder. An example of such a scheme may be

10% discount from basic premium if no claim in preceding year
20% discount from basic premium if no claim in preceding 2 years
30% discount from basic premium if no claim in preceding 3 years
40% discount from basic premium if no claim in preceding 4 years
50% discount from basic premium if no claim in preceding 5 or more years

if one claim in the year, the insured moves back one place in the scale
if two or more claims in the year, the insured moves back to the basic premium

Such schemes sometimes try to distinguish between claims that were not the fault of the policyholder and those which were. Others make no such distinction, penalising all claims according to the rules of the scheme. While many American automobile insurers do have procedures for rewarding or penalising individual experience, they usually are not as formal as the above example.

In the automobile insurance examples, claim experience determines future premiums. If the policyholder does not like the premium being charged because of his past loss experience, he can always try to find another insurer that will offer a more favorable contract. In other lines of insurance, such as worker's compensation and general liability insurance, the claim experience for the year is used to adjust the premium paid for the year. This sounds a little strange; premiums are paid in advance and, of course, the loss experience is not known until after the policy year. Such contracts contain a provision for a retroactive (retrospective) adjustment of the premium. The policyholder can be levied an additional charge or given a rebate depending on the loss experience for the policy.

## DISCOUNTED CASH FLOW ANALYSIS AND INSURANCE PRICING

### (a).The Competitive Price for an Insurance Contract

The profit maximising price for the insurance company has been presented and we have discussed how mutuals might approach the pricing problem. There is another price concept that is extremely important in insurance markets, the competitive price. There is considerable debate about whether the insurance market is competitive.[5] If we accept that it is, then it would follow that firms would not consistently make large profits. On average they would make sufficient profit to keep them in business and provide a reasonable return for the equityholders. If firms made excessive profits, new entrants would come in seeking their place in the sun. But the new competition would tend to drive down prices. On the other hand, if prices did not offer at least a reasonable rate of return to shareholders, firms would leave the industry. The resulting shortage of supply would tend to drive prices upwards. Thus prices would tend to settle at a level that promised sufficient, but not excessive, profits to existing firms. Such prices are competitive prices.

Many would disagree that insurance markets are competitive. Amongst this group are the advocates of price regulation. But even for this group, the

----------

5. see Joskow [1972] and Harrington [1984]

idea of a competitive price is important. If market forces fail to bring competitive prices then, the argument goes, regulators might. Most regulatory laws aim to set prices at the levels that would prevail under competition. The competitive price, thereby, becomes the standard for the regulator. If the regulator is empowered to set prices he(she) may estimate what price would have prevailed under competitive circumstances.

There is a third reason to look at the competitive price. If an insurance company's managers are thinking of entering a new market, or leaving an existing insurance market, they need some idea of how profitable that market is likely to be. Clearly if they can expect no profit, there is little point in being there. Conversely, with excessive profits, there is every point in being there. Much depends on the level of prices that can be charged. The competitive price is the "bottom line". If the firm is unable to sell reasonable volume at this price then it will be unable to offer its owners an adequate return on their investment. This is not to say that the firm may want to charge this price. It may be able to sell many policies at a higher price and make super profits. But the competitive price does become a guideline.

What are the features of the competitive price?

the competitive price is that which would prevail under highly competitive market circumstances

the competitive price is sufficient to deliver a positive, but not excessive, profit to the firm

the competitive price offers the shareholders (equityholders) an adequate, but not excessive, expected rate of return on their equity investment

the competitive price is sufficient to maintain the existing capital base for the insurance industry but it is not sufficient to attract new capital

the competitive price is usually the target price for regulators

the competitive price is the bottom line for a stock firm when pricing its policies

the competitive price sets the guideline for a firm in deciding whether to enter into,or exit from, an insurance market.

The competitive price, it seems, is central to our understanding of insurance markets! How can we determine the competitive price?

Up to this point, we have glossed over two of the most important and interesting issues that arise when setting prices for insurance policies. First, the payout under the insurance policy is not known when the policy is priced. The whole purpose for the policy is that the loss is uncertain. How can our pricing policy cope with this uncertainty? Second, any losses that may be paid under the policy will be paid in the future but the premium is paid in advance. How can our pricing policy cope with cash flows that arise

at different points in time? These two issues are central in determining the competitive price. The solutions lies in Discounted Cash Flow Analysis.

(b) Discounted Cash Flow Analysis (DCF)

It is generally believed that most people have an inate preference for the present over the future. We are myopic. If offered one dollar today or the promise (with certainty) of one dollar next year, we generally prefer the dollar today. Since money indicates purchasing power, this implies that present consumption is more valuable than future consumption. This idea is known as time preference. Time preference explains why interest rates are positive. If a person is to be induced to lend his money, he must be offered some compensation to overcome this time preference. That compensation is interest. Time preference also explains why the borrower is willing to pay interest. The borrower gains use of money now but has to repay in the future. Time preference indicates that he would be willing to give up more than a dollar in the future to gain access to today's dollar.

(b).i. The Net Present Value.

Given pervasive time preference, cash flows arising at different points in time have different value. The earlier a cash flow, the greater its value. The interest rate therefore becomes a rate of exchange between cash flows at different points in time. Just as we have rates of exchange to convert British pounds and Japanese yen into U.S. dollars, so too we have rates of exchange to convert dollars that might arise in one or two years time into today's dollars. This rate of exchange is the interest rate or, more generally, the discount rate. If we had a basket of foreign currencies, we would choose to convert it all to one currency (perhaps dollars) to ascertain its value. So with time preference we convert all cash flows to a single point in time (usually the present) in order to ascertain their total value. Thus we often talk of the net present value of a set of cash flows.

********************************ILLUSTRATION***********************************

If you have $1 now and you are promised $1 in one year, what is the total value of your wealth. To solve this, we need to know the discount rate. Time preference, we assume, is measured by a discount rate of 10%. All the cash flows must now be converted to the same time, e.g. the present. The $1 you already have is already a "present dollar" and needs no conversion. But the promised dollar comes in a year and needs to be discounted to give a present value

net
present value = (one dollar) + (one dollar discounted one year)
(NPV)

$$= \$1 + \frac{\$1}{(1 + 0.1)}$$

$$= \$1 + \$0.91$$

$$= \$1.91$$

The same technique can be used to discount more distant cash flows. Thus with \$1 now and \$1 for each of the next three years, the present value is

$$NPV = \$1 + \frac{\$1}{(1+0.1)} \quad \frac{\$1}{(1+0.1)^2} \quad \frac{\$1}{(1+0.1)^3}$$

The method also can be used with negative cash flows. Consider the following set of cash flows to be discounted at 10

| Time | Cash Flow $C^i$ | Present Value $C^i/(1+k)^i$ | | |
|---|---|---|---|---|
| 0 | 200 | | = | 200.00 |
| 1 | 1000 | $1000/(1+0.1)$ | = | 909.09 |
| 2 | 200 | $200/(1+0.1)^2$ | = | 165.29 |
| 3 | -500 | $-500/(1+0.1)^3$ | = | -375.66 |
| 4 | -400 | $-400/(1+0.1)^4$ | = | -273.21 |
| | net present value | | | 625.51 |

********************************************************************************

Thus, to calculate the present value, we sum the discounted values of the cash flows for each year. This is expressed mathematically as follows. $C^i$ refers to the cash flow in year "i" and k is the discount rate

$$NPV = C_0 + C_1/(1+k) + C_2/(1+k)^2 + C_3/(1+k)^3 \ldots\ldots\ldots\ldots \text{etc.}$$

$$= \Sigma C_i/(1+k)^i$$

(b).ii. The Internal Rate of Return.[6].

Discounted cash flow analysis may be used to calculate the <u>net present value</u> of a set of cash flows or the <u>internal rate of return</u> on those cash flows. The internal rate of return is an implicit discount rate earned on the cash flows. Technically it is defined as the discount rate that equates the

----------

6. Those wishing to use this technique extensively should note that problems sometimes arise such as multiple internal rates of return. These issues are dealt with in most basic capital budgeting textbooks

cash flows to a zero net present value.

******************************ILLUSTRATION**************************************

Consider the following set of cash flows. The net present value is calculated using various discount rates.

| time | cash flow | present value at discount rate: 0% | 5% | 10% | 15% | 20% |
|---|---|---|---|---|---|---|
| 0 | -1243 | -1243 | -1243 | -1243 | -1243 | -1243 |
| 1 | 500 | 500 | 476.2 | 454.5 | 434.8 | 416.7 |
| 2 | 500 | 500 | 453.5 | 413.2 | 378.1 | 347.2 |
| 3 | 500 | 500 | 431.2 | 375.7 | 328.8 | 289.4 |
| net present value (rounded) | | 257 | 118 | 0 | -101 | -190 |

As the discount rate increase, the value of the future (positive) cash flows falls. The current negative cash flow of $1243 is, of course, unaffected by the discount rate. Consequently, the net present value of the stream of cash flows declines as the discount rate increases. At a discount rate of 10%, the net present value is zero. Thus 10% is the <u>internal rate of return</u>. The set of cash flows outlined is equivalent to investing funds at 10%. To see this consider making an investment of $1243. Would this exactly sustain three successive annual income instalments of $500. Yes it would. Consider the following balances,

| Year | beginning of year | end of year | |
|---|---|---|---|
| 0 | invest 1243 | (1243) times (1+0.1) | = 1367.3 |
| | | less 500 | = 867.3 |
| 1 | balance 867.3 | (867.3) times (1+0.1) | = 954.03 |
| | | less 500 | = 454.03 |
| 2 | balance 454.03 | (454.03) times (1+0.1) | = 499.5 |
| | | less 500 | = zero (approximately) |

Subject to the rounding error, the investment does sustain annual payments of $500 if the interest rate is 10%. In other words, we may think of the internal rate of return as the implicit interest earned on the set of cash flows.

*******************************************************************************

(b).iii. Risk Adjusted Discount Rates

Having coped with the timing of cash flows, we must now address risk and uncertainty. Future cash flows are rarely, if ever, certain. Insurance cash flows, in particular, are highly uncertain. Most policies pay nothing to the policyholder simply because no loss is sustained. But, if a loss is sustained, it may be minor or serious. Thus there is a whole range of possible values which might be paid under the insurance policy and neither the insurer not he policyholder know in advance what the payment will turn out to be. There is an exception to this statement. A life insurance policy might be held until the eventual, and inevitable, death of the policyholder. But even in this example, the timing of the payment is not known.

In Chapter 3, the demand for insurance was explained by assuming the people were generally averse to risk. If presented with a certain payment of $100 in one year's time or a 50% chance of $200, payable at the same time, the risk averse person will go for the certain $1. From the earlier discussion, the risk averse person prefers a certain outcome to an uncertain outcome having the same expected value. This observation provides the clue for dealing with risk when valuing the cash flow. If we were to discount the expected value of a cash flow to give a present value, we would overvalue the cash flow (assuming it to be a positive cash flow). Replacing the cash flow by its expected value is to treat an uncertain cash flow as though it were certain. But an adjustment can be made in the discount rate to reflect the uncertainty.

**********************************ILLUSTRATION*************************************

Consider the folowing example;

The present value of $100 for certain in one year, given a discount rate of 10% is

PV = $100/(1+0.1) = $90.91

Now consider a gamble involving a 50% chance of $200 to be paid in one year. The expected value is $100 but we know that it is worth less than the $100 for certain. The distaste for risk can be reflected in the discount rate. Suppose that the discount rate is increased to 13%. The present value, <u>adjusting for risk</u>, is;

PV = $100/(1+0.13) = $88.50

The new discount rate is the <u>risk adjusted discount rate</u>

*************************************************************************************

This is a standard method of dealing with risk. The expected values of the cash flows are used, but they are discounted using risk adjusted discount rates. Just how the adjustment for risk is made will be considered in later chapters.

ESTIMATING THE COMPETITIVE PRICES FOR INSURANCE CONTRACTS

(a). Insurance Cash Flows

(i) Premiums and Expenses. Most insurance policies require the payment of the insurance premium "up front." Payment is normally required for the policy to come into effect. The insurance firm therefore has money on hand to pay for up front expenses such as marketing and underwriting costs and the payment of commissions to agents and the establishment and keeping of records. After meeting these expenses the insurer then establishes a reserve from which future claims payment can be met. But the reserve is an internal accounting device used by the insurer; it is not a cash flow. We are only concerned with cash flows into, and out of, the firm.

Sometimes, one encounters exceptions to the prepayment of premiums. For example, premiums may sometimes be payable in instalments (e.g., on personal automobile policies). Another exception is the retroactive rating often encountered in liability insurance contracts such as workers compensation insurance. With these contracts, a premium adjustment is made after the contract has expired. The adjustment depends on the loss experience of the individual policy or of the group in which it is classified.

(ii). Claim Payments and Claim Settlement Expenses. The remaining major cash flow is the loss payment. For the individual policy the loss payment is uncertain both in timing and usually in amount. For example, with a policy covering my house, I may have a loss any time within the year during which cover is contracted and the loss, if it arises, may be a small or large value. For the insurance firm which has a large number of policies, the effects of diversification brings more predictability in the aggregate of lost payments. Thus if the portfolio is large enough and if losses have a low correlation, the insurer will have a steady and predictable outflow of loss payment to make.

In addition to paying the claim, the insurer usually incurs expenses. Often the loss must be investigated and negotiated. Sometimes, the insurer must pay salvage expenses or lawyers fees to settle a third party liability claim. The settlement expenses are usually incurred in the time leading up to settlement and depend, to some extent, on the occurence and size of the claim.

(iii). Other Cash Flows. Probably the major cash flow not covered is taxation. In writing policies, the insurer incurs tax liabilities which also must be anticipated. There may be other minor cash flows that arise in particular cases. But only if the items in question are cash flows is it proper to use them in discounted cash flow analysis. We have mentioned that reserves are not cash flows and should be counted. Similarly, depreciation is a similar accounting concept that does not correspond to cash flowing into, or out of, the firm. Discounted cash flow analysis means what it says! Only cash flows that enter or exit the firm should be accounted. Just because an accounting entry appears somewhere does not necessarily mean that there is a cash flow.

(b). A Simple Example of Premium Calculation Using Discounted Cash Flow

For the simple example, consider a homeowners policy. The policy runs for one year, thus losses may arise on any one of the 365 days of cover. As a first approximation, let us assume that a loss is equally likely one each of the 365 days and further simplify things by suggesting that, on average, the loss will arise six months after inception of the contract. Given a further delay of say two months for the filing and settlement of the insurance claims, it transpires that the insurer holds the typical dollar of premium (net of expenses) for eight months before paying the loss. The expected value of loss under the policy is $200 and the monthly interest rate is 0.5%. Expenses, which are incurred at inception, are $30. Simple use of DCF establishes the present value of the contract.

$$PV = (-30) + \frac{(-200)}{(1 + 0.005)^8} = \$222.18$$

Discounted cash flow analysis tells us what a set of cash flow is "worth" as a present value. Well, this set of cash flows is worth $222.18. Consequently, this would be the competitive premium.

Another way of looking at this problem is to use the alternative discounted cash flow tool; the internal rate of return. Recalling that the internal rate of return is the implicit discount rate that yields a zero net present value for a set of cash flows, consider the internal rate of return on the following;

$$0 = (\$222.18) + (-30) + \frac{(-200)}{(1 + r)^8}$$

Where "r" is the internal rate of return. Of course, these are the cash flows used in the previous example to calculate the competitive premium. If this problem is solved for the discount rate "r", the answer, not surprisingly, turns out to be a monthly rate of 0.05%. There is a clear symmetry between the NPV and the IRR. If the premium is set equal to NPV, then the insurance contract bears an implicit interest rate (or IRR) which is equal to the discount rate used in calculating the NPV. This discount rate is referred to as the "cost of capital" under the insurance contract. The insurer has use of the policyholders' funds for eight months and the cost of using these funds is 0.5% per month; thus the term "cost of (insurance) capital." The cost of capital will become important later when considering issues such as the optimal leverage of an insurance firm.

The same basic ideas can be used to estimate competitive premiums for more complex insurance policies. Allowance can be made for the prospect that losses are not paid on a preassigned date, but may be paid at any future date. The method also can be used for insurance policies in which payments may be made many years into the future, such as life insurance or medical malpractice.

X (c) <u>Frequent Payout</u>

A more useful way of describing the prospective losses on the policy described above recognizes that losses may arise early or late in the policy life. Thus, instead of attributing the expected loss of $200 to one point in

time, (i.e. six months into the policy life plus two months settelment delay), it may arise in any of the twelve months for which the contract runs. Recognizing the two month delay in filing and settling claims, and assuming that the probability of loss is equal in each month, then the expected value of payout in the third month (loss arises in month one but is settled two months later) is \$200 divided by twelve = \$16.7; and in the fourth month it is \$16.7; and so on. The expected cash flows on the policy may now be presented as follows

| month | cash flow |
|---|---|
| 1 | -30 |
| 2 | 0 |
| 3 | -16.7 |
| 4 | -16.7 |
| 5 | -16.7 |
| 6 | -16.7 |
| 7 | -16.7 |
| 8 | -16.7 |
| 9 | -16.7 |
| 10 | -16.7 |
| 11 | -16.7 |
| 12 | -16.7 |
| 13 | -16.7 |
| 14 | -16.7 |

The competitive premium now may be calculated as the present value of this set of cash flows. For convenience, we assume that the cash flow in each month takes place on the first day of that month.

$$\text{Competitive Premium} = 30 + 0/(1 + .005)^1 + 16.7/(1 + .005)^2 + 16.7/(1 + .005)^3 + \dots\dots\dots + 16.7/(1 + .005)^{14}.$$

$$= \$223.07$$

The difference between the premium calculated here (\$223.07) and that calculated earlier (\$222.2) is small, but it is instructive to know why there is any difference at all. The difference may be accounted for by two factors. In the first place there is a rounding error. The expected loss of \$200 was divided by twelve and then simply rounded to give \$16.7. Greater accuracy could be obtained by using more decimal places. More substantially, the second calculation gives a more accurate picture of the timing of the cash flows. Given non zero discount rates, one dollar payable after eight months does not have the same present value as equal payments of 1/12 of a dollar spread over months three to fourteen. To see this second point more clearly, simply compare the payment of \$2 at the end of one year with the payment of \$1 immediately and \$1 after two years. With a discount rate of 10%, the respective NPV's are \$1.81818 and \$1.82645. Thus it should be clear that the more accurately we represent the payout pattern over time, the more accurate the premium calculation will be. So our example would have yielded greater accuracy with weekly or even daily payout data. But this additional accuracy

would be achieved at a cost in terms of greater complexity and in terms of the higher costs of frequent data collection. Thus premium calculation requires that we trade off accuracy with cost and complexity when settling on the timing of the policy cash flows.

Normally, an insurance firm will not set a premium separately for each individual policy but will set premium rates collectively for groups of policies that are similar in terms of their loss characteristics. For example, the premium rates will be set collectively for the population of insureds of a given age and sex. In such a setting, the insurer must set the premium such that total premiums are sufficient to cover total loss payments and give the insurer an adequate return on its investment. The insurer will have to make many loss payments on each cohort of policies. For example, some 10,000 policies written in January might be expected to give rise to claim payments in February, March, April, and so on. Thus we can set the premium by estimating the expected claim payments month by month into the future. Consider, for example, that from the 10,000 policies written in January the insurer expects to be settling fifty claims each month for the next fifteen months. The settlements run over year because of settlement delays. On average, each settlement is expected to cost $500. With expenses at $10 per policy, the NPV of the cash flows is

$$NPV = (10{,}000 \times -\$10) + \sum_{1}^{14} \frac{50 \times -\$500}{(1 + .005)^{i}}$$

$$= -\$460{,}415$$

This layout assumes that loss settlements arise on the first day of February and monthly thereafter for the next fifteen months. If the policies are assumed to be identical, the insurer can recover the total of $460,415 in equal premiums per policy of $46.04. The competitive level of premium is therefore $46.04 per policy.

## LIFE INSURANCE PREMIUM CALCULATION

(a) <u>Term Life Insurance - Single Premium</u>

The basic ideas for calculating the competitive premium for life insurance and for non life insurance are identical. Each application simply requires discounting of the expected cash flows at an appropriate discount rate. Differences in application arise largely because non life contracts usually run for one year or less but most life insurance contracts remain in force for many years and require payments of successive annual premiums. Calculations over many years do get more complex. But the good news with life insurance is that the calculation of the expected values of loss payouts is usually much easier since the value to be paid in the event of loss is agreed contractually (often at a fixed sum). Moreover, estimation of the probability of loss usually is easily extracted from mortality tables.

The two basic forms of life insurance are term insurance and whole life. Other forms that you may have heard of, such as Universal Life, are usually one or other of these two types combined in a financial package with some investment vehicle such as a mutual fund. Term insurance is simple insurance protection for a stated period, say one year. If the insured dies in the period, the loss is settled and the policy expires. If the insured survives, no loss is paid and the policy expires at the end of the period. In contrast, whole life offers continuous protection up to the inevitable demise of the insured whenever that might eventually occur. Unless the policy is cancelled, say by non payment of premiums, ultimate payment of the policy money is certain. The only uncertainty is when the insured will die and thus when the claim will become payable. Usually, whole life premiums are constant year by year in spite of the fact that mortality risk usually increases with age.

To proceed with our excursion into life insurance rating it is useful to have available a mortality table. In the Appendix, the CSO, 1980 Standard Mortality Tables are reproduced. These will be the required basis for life insurance rating by 1989. The features of the Table will become apparent in our rating examples.

Consider a three year, single premium term insurance with a <u>face</u> value of $1,000 for a male age thirty. Under such a contract, the policyholder will

pay a single premium up front and will receive three years life insurance protection in return. If he should die within the three years, the policy will pay \$1,000. The Mortality Table shows that the probability of death in the first year for a thirty year old male is 0.00173. If he survives the first year (for which the probability is 1.0 - 0.00173 = 0.99827) then there is a probability of 0.00178 of his dying in the second year. Similarly, if he survives the first two years (probability = (1.0 - 0.00173) x (1.0 - 0.00178) = 0.99649), then there is a probability of 0.00183 that he will die in the third year. In rating the policy, we can use these probabilities directly to calculate the expected cash flows under the policy.

The task now is to estimate the expected values of cash flows under the policy. First, we will make things simple by assuming that, if the insured should die in any year, he will die on the last day of the year (New Year parties are particularly violent). A second useful thing to remember is that people only die once. Thus, if the policyholder is to die in the second year, he must first survive the first year of the policy. Thus, to find the probability of death of the policyholder in the second year of the policy we must multiply the probability that he will survive the first year (0.99827) by the probability of death in the second year assuming that he has already lived through the first year of the policy (0.00178). Similarly, the policyholder has to survive the first two years (probability 0.99649) to have the dubious priviledge of facing death in the third year. Using these ideas, the NPV of the policy is

$$NPV = \frac{(0.00173)(-\$1,000)}{(1 + 0.1)} + \frac{(0.99827)(0.00178)(-\$1,000)}{(1 + 0.1)^2}$$

$$+ \frac{(0.99827)(0.99822)(0.00183)(-\$1,000)}{(1 + 0.1)^3}$$

$$= -\$4.41$$

Thus the competitive premium for this three year single premium term policy is \$4.41 payable at the beginning of the contract.

(b) <u>Term Life Insurance - Annual Premiums</u>

Life insurance policies come in all shapes and sizes. As in any other product market, individual producers may be able to secure market advantage by differentiating their product from those of other producers or by offering a range of products each aimed at buyers of different profile. Thus it is difficult to say what is a standard term life policy; each has different bells and whistles attached. But our concern here is with the basic principles of pricing, not with showing every conceivable pricing formula for every conceivable policy. If the basic priciples are laid down, they should be adaptable to most product innovations. Thus we will continue with another fairly basic idea, that of annual premiums.

A quick glance at the mortality tables will reveal, that after the initial surge of infant mortality, death rates tend to increase with age, accelarating in pace from middle age. Consider the implications of this pattern if the three year term insurance descibed above was to be financed, not with a single premium, but with three annual premiums payable at the beginning of each of the policy years. The simplest way to achieve this is to calculate each annual premium from the mortality rate for that year. Given the mortality rate between ages 30 and 33, the respective annual premiums would be (remember that premiums are payable at the beginning of each year whereas losses are assumed to be payable at the end of each year);

$$\begin{array}{ll} \text{Year 1} & (0.00173)(\$1{,}000)/(1 + 0.1) = \$1.57 \\ \text{Year 2} & (0.00178)(\$1{,}000)/(1 + 0.1) = \$1.62 \\ \text{Year 3} & (0.00183)(\$1{,}000)/(1 + 0.1) = \$1.66 \end{array}$$

The quick witted may have anticipated that there should be some relationship between these annual premiums and the single premium calculated earlier. It might seem that the NPV of these annual premiums should be identical to the single premium. This turns out to be the case. The NPV of the annual premiums is;

$$\begin{aligned} NPV &= \$1.57 + \$1.62/(1+0.1) + \$1.66/(1+0.1)^2 \\ &= \$4.41 \end{aligned}$$

which is the same as the single premium payable immediately.

(c) Level Premiums

In the previous example with annual premiums, the premiums increased as the insured became older and the mortality rate correspondingly increased. If this form of term insurance were the only form of life insurance available, it is clear that we all would be faced with prohibitively high premiums as we became old and, perhaps, most in need of the insurance protection. When we most needed protection we would be least able to afford it. Another type of insurance policy addresses this issue; the level premium contract. Level premiums can be illustrated with a simple term contract but they become central to the understanding of whole life contracts. For illustration we keep with the thirty year old male seeking three years protection, but now the task is to calculate an annual premium payable at the beginning of each of the three years. The catch is that the premium must be the same value in each of the three years. The calculation must also recognize that the policyholder may not survive to pay later premiums i.e. if the policyholder dies in the first year he will not pay the second and subsequent premiums.

The idea is quite simple though the actual calculation can become messy. We must set a constant premium, call it P, for which the expected NPV is equal to the single premium, bearing in mind the probabilities that the insured will survive to pay each of the respective premiums. Using the notation $S_{i,j}$ for the probability that the insured will survive from age "i" to age "j", the appropriate formula is;

-{NPV of the expected cash flows (which is equal to the single premium)}
– NPV of the expected level annual premium premium payments.

The NPV of the level annual premiums is:

$$\text{NPV(pemiums)} = P + \frac{P\ S_{30,31}}{(1 + r)} + \frac{P\ S_{30,32}}{(1 + r)^2}$$

$$= -(\text{NPV(cash flows)})$$

which can be solved for P as follows

$$P = \frac{-(\text{NPV(cash flows)})}{1 + \dfrac{S_{30,31}}{(1 + r)} + \dfrac{S_{30,32}}{(1 + r)^2}}$$

The simplest and most accurate way to obtain any survival probability, $S_{i,j}$ is to divide the number alive in year j by the number alive in year i. For a thirty year old male, the chances of surviving to ages 31 and 32 can be calculated using the CSO 1980 Mortality Tables;

$$S_{30,31} = 9{,}563{,}425/9{,}579{,}998 = 0.99827$$

$$S_{30,32} = 9{,}546{,}402/9{,}579{,}998 = 0.9964931$$

Returning to the problem of a three year term policy for a thirty year old male, we can now calculate the level annual premium with the above formula;

$$P = \frac{4.41}{1 + \dfrac{0.99827}{(1 + 0.1)} + \dfrac{0.9964931}{(1 + 0.1)^2}}$$

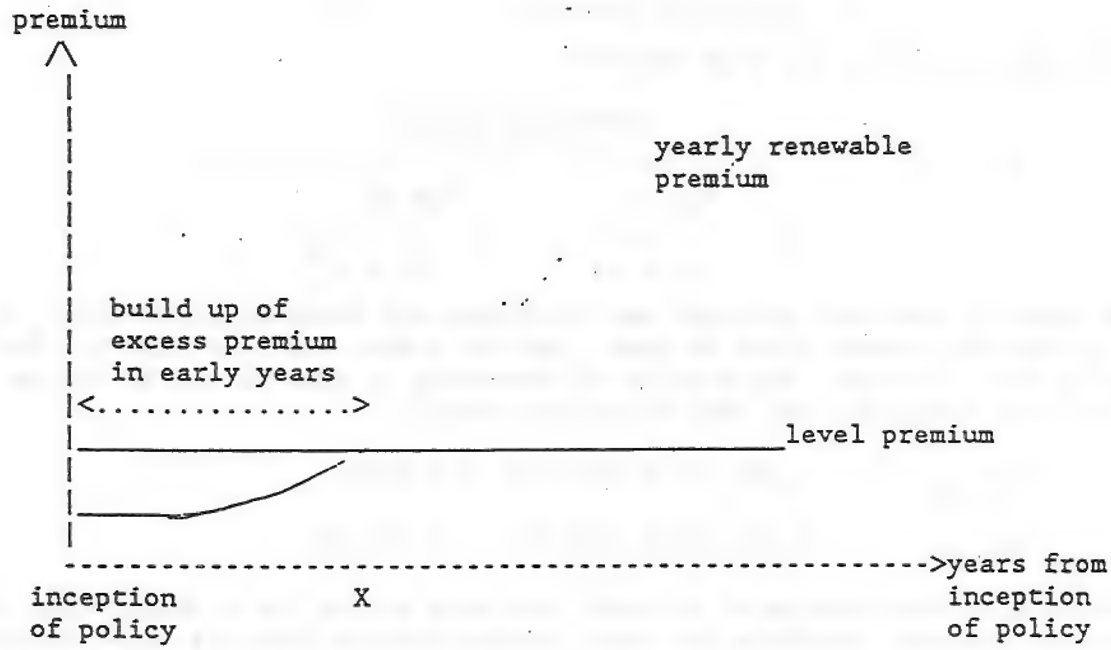
$$= \$1.615$$

Comparison of the level annual premium with the earlier varying premiums example, it will be seen that the level premium is higher than the one year premium during the first year and lower than the one year premium during the third year. During the second year, the one year premium and the level premium are somewhat similar. What happens is that the level premium causes the early premiums to subsidise the later premiums. In early years, the premium is more than sufficient to reflect the mortality rate. The surplus in these years, together with any investment income from investment of the surplus, can be used to keep the later premiums low despite the high mortality rates in the final years of the policy.The early surplus premium together with investment income subsidizes the deficient later premiums.

In our three year example, the value of the subsidy is small since the period is short and there is little difference in the mortality rates between ages 30 and 33. But if the policy were to continue for, say 20 or 30 years at

years which could be used to keep the later premiums relatively low in spite of the high mortality incidence during those later years. The surplus which is built up under the level premium principle is known as the <u>legal reserve</u>. The idea is illustrated in Figure 4

Figure 4




The legal surplus built up under level premium policies carries some implications for the design of policies. Clearly it would be silly if insurance compnies could cancel policies whenever they chose. If so the insurance company would simply allow the policyholder to pay a level premium for a number of years then simply cancel the policies when the surplus was at a maximum. This would be tantamount to theft of the reserve from the policyholder. Similarly, the policyholder would be severely penalised if he or she were to cancel the policy in the early years since the subsidy would be lost. For these reasons longer period level premium policies, particularly whole life policies to be discussed next, contain surrender value provisions. Should the policyholder choose to cancel he or she may be able to recover part of the accumulated reserve. The amount that may be recovered is known as the surrender value. But this is digressing a little from the calculation of premiums.

(d) <u>Whole Life Insurance Premiums</u>

In describing a term policy with a level premium and in discussing a very long term for cover, we have come very close to describing a whole life policy. If a whole life policy was undertaken for a long enough period at a level premium, the policyholder is certain to die during the period of cover. Assuming the policyholder choses to renew the policy each year, the prospect that the insurer will pay a claim is certain, the only doubt is when. Such a policy is a <u>whole life policy</u>. Thus, in effect, the whole life policy is an

early subsidy and typically, the insured has the right to cancel the policy and to recover the surrender value. In practice, there are contractual rules for deciding how much should be recovered and these will take account of the insurer's expenses (which are typically loaded up front) and of investment earnings.

In trying to determine the value of a whole life policy, and therefore in determining its premium, one problem encountered is in determining a boundary on the life of the policy. If the policy is to remain in force over the life of the insured, we must discount the expected cash flows over the prospective lifetime, i.e. we must determine the expected cash flow for each year in the future. At what point do we stop. Clearly there is no point in looking at the expected cash flow when the insured reaches the age 200 since the probability of reaching that age is as close to zero as we may imagine. More reasonably, do we stop at age 105, or at 110, or where. Theoretically, we should stop at the age for which the probability of living an extra year is zero. Of course, we do not know this age. Even if there were no documented evidence that anyone had survived beyond say 110, this would not imply a zero probability for someone breaking the record. This problem has been solved in a quaint and convenient, though somewhat arbitrary, fashion.

On examining the Commissioners' Mortality Tables, it appears that no-one lives beyond the age 100. According to the Tables the probability of death in one's 100th year (for both males and females) is 100%. Of course this is untrue. This is not a creation of that sinister, but fictitious, club the "Actuaries Euthanasia Society". Rather, it does serve as a convenient approximation given the very low probabilities of surviving to ages beyond 100. Thus survival probabilities beyond this age are approximated at zero. A nice corollary which is consistent with the use of such tables, is that the whole life insurance policy will pay the face value at age 100 even if the insured is still alive.

The convenience of terminating the Mortality Tables at 100 permits the NPV to be calculated on cash flows up to that age, bearing in mind a) that those surviving to 100 receive the face value, and b) that a level premium will be paid by those surviving to each future age. With these thoughts in ming, we can use the previous analysis for determining the level premiu to be charged for a whole life policy. We simply apply the previous practice of setting the NPV of expected claim payment equal the NPV of future level premium payments.

$$\text{NPV(cash flows)} = \sum_{i=1}^{100-t} \frac{m_{t,t+i}(FV)}{(1+r)^i}$$

and

$$P = \frac{\qquad}{1 + \sum_{i=1}^{100-t} \frac{S_{t,t+i}}{(1 + r)^{i}}}$$

where $t$ is the current age of the insured $m_{t,t+i}$ is the probability that an insured currently of age $t$ will die at age $t+i$ NPV is the present value of claim payment $P$ is the level premium $S_{t,t+i}$ is the probability that the insured currently of age $t$ will survive to age $t+i$ $r$ is the discount rate

This premium formulation is pretty basic. As mentioned earlier, many policies are designed with individual features. For example, some policies only have a limited number of premium payments; others have dividend payments which depend on investment income; etc. Clearly accommodations will need to be made in the premium calculation to suit the individual policy. Moreover, we have not addressed the policy expenses in order to simplify our exposition. The expense cash flows also will have to be added in to derive the NPV of payouts under the policy.

9



# CHAPTER 14

## RETURN TO INSURANCE PRICING

Chapter 12 examined the financial structure of an insurance company. This was achieved by tracing the cash flows of the insurer and relating these in a simple algebraic model. With this model we were able to identify how much profit the insurer made and what rate of return it was able to deliver to its equityholders. The model was then used to see whether the insurer would be able to meet its obligations or whether it might become insolvent. The same cash flow model is now used for a different purpose. In this Chapter, we return to the subject of setting premiums.

Chapter 4 set the stage for looking at insurance pricing. There insurance prices for individual policies were set simply by discounting the expected cash flows from the insurance contract at some appropriate discount rate. We will now consider how the setting of premiums fits into the overall financial famework of the insurer. This will reveal whether any given premium level will result in an adequate return on equity. The financial framework reveals the effects of leverage and investment income on the selection of premium rates. We will also be able to select the appropriate discount rate to use in the discounted cash flow analysis. The chapter will go on to examine the effects of inflation on premium setting. Finally, the chapter will examine a rather curious phenomenon. Insurance prices appear to fluctuate in a cyclical pattern. We will see whether the cash flow model can explain this mystery.

### A REVIEW OF CHAPTER 4

The general principles of insurance pricing were laid out in Chapter 4. Competitive pressures would lead towards the classification of risks according to their expected losses and/or expenses. Within any risk classification, the competitive price could be established simply by discounting the cash flows that arise for the insurance company when it iss s its polcies. The formula for the competitive price was;

$$\text{Price} = \text{NPV} = \Sigma\, C_i/(1+k)^i \qquad (14.1)$$

where $C_i$ is the cash flow (loss/expense/tax) at time "i"
$k^i$ is the risk adjusted discount rate

This formula was intended to achieve a competitive expected rate of return on equity but, in fact, explicit attention was never given to equity. The implied reasoning used there was that all cash flows have an economic value which is measured by the usual discounting formula. By trading at this value, the insurance company was making a normal economic return. This would be sufficient to maintain the supply of equity to the firm since it would offer investors a competitive expected return on their investment. But it would not be sufficient to yield excessive expected profits. This is not necessarilly the profit maximising price but it does show the "bottom line". If the firm is unable to maintain this price in the long run, it should close its doors.

We return to these ideas but now we may focus explicitly on the effects of pricing on the return on equity and we may see the effects of investment income.

U.S. CLAIMS COSTS INDEXES VS.CONSUMER PRICE INDEX(CPI)
(FROM 1967 TO 1986)
(1967=100)

| YEAR | TOTAL AUTO | AUTO BODI. | AUTO PROP. | AUTO PHYS. | W.C. | OTHER BODI. | OTHER PROP. | GLASS |
|---|---|---|---|---|---|---|---|---|
| 1967 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |
| 1968 | 110.8 | 108.5 | 112.4 | 113.5 | 109.6 | 108.2 | 107.9 | 105.6 |
| 1969 | 119.3 | 118.2 | 120.1 | 120.5 | 119.6 | 117.8 | 115.8 | 113.3 |
| 1970 | 126.5 | 128.3 | 125.5 | 124.6 | 130.9 | 128.3 | 122.5 | 120.5 |
| 1971 | 139.5 | 139.4 | 139.3 | 139.0 | 143.8 | 139.9 | 133.2 | 129.6 |
| 1972 | 149.6 | 147.1 | 150.9 | 152.2 | 155.0 | 147.7 | 142.1 | 135.2 |
| 1973 | 160.7 | 157.1 | 163.3 | 164.5 | 171.4 | 157.2 | 153.7 | 142.8 |
| 1974 | 171.1 | 171.6 | 172.7 | 172.3 | 189.0 | 171.8 | 161.9 | 149.5 |
| 1975 | 190.5 | 191.9 | 190.1 | 189.1 | 218.4 | 192.2 | 177.0 | 159.2 |
| 1976 | 212.1 | 214.5 | 210.3 | 210.2 | 244.7 | 214.8 | 197.2 | 175.6 |
| 1977 | 231.8 | 236.2 | 228.7 | 228.1 | 275.4 | 235.8 | 214.3 | 185.2 |
| 1978 | 251.8 | 258.5 | 249.6 | 244.3 | 304.1 | 256.9 | 232.7 | 202.2 |
| 1979 | 275.2 | 284.6 | 272.3 | 265.5 | 336.6 | 282.1 | 252.7 | 216.1 |
| 1980 | 306.9 | 316.3 | 304.3 | 297.8 | 368.5 | 314.8 | 271.3 | 234.1 |
| 1981 | 348.1 | 355.5 | 345.3 | 340.9 | 412.2 | 352.9 | 295.8 | 257.4 |
| 1982 | 386.3 | 400.9 | 381.2 | 371.4 | 462.8 | 399.3 | 322.1 | 275.6 |
| 1983 | 413.4 | 437.3 | 404.4 | 388.3 | 504.8 | 432.7 | 347.0 | 290.0 |
| 1984 | 434.2 | 472.6 | 419.4 | 394.3 | 551.9 | 457.8 | 367.4 | 299.7 |
| 1985 | 456.6 | 507.1 | 431.1 | 395.6 | 588.0 | 495.8 | 395.1 | 308.2 |
| 1986* | 480.7 | 538.3 | 451.8 | 410.9 | 615.4 | 535.8 | 406.5 | 320.9 |

| YEAR | THEFT | BOILER & MACH | FIRE | HOME OWNER | COMMER PERIL | INLAND MARINE | P&L TOTAL | CPI |
|---|---|---|---|---|---|---|---|---|
| 1967 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |
| 1968 | 106.9 | 105.2 | 106.3 | 107.0 | 106.4 | 106.6 | 109.6 | 104.2 |
| 1969 | 113.2 | 111.0 | 115.6 | 115.8 | 115.9 | 112.7 | 118.1 | 109.8 |
| 1970 | 119.0 | 116.8 | 123.8 | 123.4 | 125.0 | 118.4 | 126.2 | 116.3 |
| 1971 | 127.0 | 123.3 | 135.5 | 133.6 | 137.4 | 125.7 | 138.4 | 121.3 |
| 1972 | 132.9 | 131.9 | 146.0 | 144.0 | 147.0 | 132.1 | 148.5 | 125.3 |
| 1973 | 146.4 | 140.5 | 159.3 | 156.8 | 160.8 | 143.1 | 163.7 | 133.1 |
| 1974 | 155.5 | 153.9 | 172.4 | 170.4 | 172.6 | 153.4 | 173.8 | 147.7 |
| 1975 | 167.3 | 166.0 | 185.5 | 183.5 | 187.4 | 166.2 | 192.5 | 161.2 |
| 1976 | 180.3 | 181.7 | 199.8 | 199.9 | 202.0 | 179.1 | 213.4 | 170.5 |
| 1977 | 198.5 | 198.9 | 216.1 | 217.0 | 216.6 | 192.8 | 235.1 | 181.5 |
| 1978 | 215.9 | 218.2 | 235.1 | 235.9 | 236.9 | 209.2 | 257.5 | 195.4 |
| 1979 | 235.9 | 235.2 | 256.8 | 257.7 | 259.5 | 227.5 | 281.8 | 217.4 |
| 1980 | 257.1 | 256.7 | 278.3 | 278.7 | 282.6 | 250.0 | 310.8 | 246.8 |
| 1981 | 276.7 | 284.2 | 298.5 | 306.1 | 304.8 | 269.8 | 345.8 | 272.4 |
| 1982 | 295.0 | 295.2 | 317.1 | 332.8 | 327.4 | 287.1 | 383.1 | 289.1 |
| 1983 | 313.7 | 312.6 | 331.2 | 354.4 | 345.5 | 301.4 | 408.8 | 298.4 |
| 1984 | 338.2 | 336.5 | 342.1 | 371.9 | 360.5 | 319.7 | 432.3 | 311.1 |
| 1985 | 361.3 | 342.7 | 349.6 | 387.5 | 371.4 | 336.3 | 456.8 | 322.2 |
| 1986* | 318.6 | 351.9 | 356.3 | 400.5 | 381.7 | 352.2 | 479.2 | |

* ESTIMATED
DATA SOURCES: MORTON E. MASTERSON, Best's Review-P&L-(Sept.1976) and Best's Insurance Management Reports(Perspectives:5/25/'81 &4/14/'86).

## THE FAIR OR COMPETITIVE PREMIUM; HISTORICAL VIEWS

In a regulatory context, the level of prices which would prevail in a competitive environment has become the main yardstick. Regulators of property-liability insurance companies, utilities, airlines and other industries have tended to view the competitive price as the ideal. But even in a non regulated environment, the competitive price is an important concept since it provides a barometer of how efficiently the market is working. The competitive price also tells the managers of an insurance firm how much they must charge in order to deliver an acceptable return on equity to their equityholders.

In regulatory circles, the competitive price is often referred to as the "fair" price. Fair is a normative concept but it conveys the idea that such a price is just sufficient to cover the costs of the insurer, including the provision for an adequate return to equityholders; but the price is not so high as to offer excessive expected return to equityholders. Such a price may be considered to be "reasonable" to consumers and to suppliers of insurance services alike. Furthermore, such a price would be just sufficient to maintain a supply a capital resources to the insurance industry

In property-liability insurance, it is usual to think of prices in terms of their implications for the expected level of underwriting profit. During the 1920,s The National Association of Insurance Commissioners recommended a standard for underwriting profit of 5%. For several decades this standard has been used as a reference point in the commercial setting of rates and in rate regulation (though subject to some small variation across lines). To our knowledge, there is no theoretical, or other objective reasoning behind the 5% standard, though it has for many years been accepted as a "reasonable" profit margin. The use of a 5% target in rate making is based on expected profits. Thus we will restate the underwriting profit as an <u>expected value</u>;

$$\text{Underwriting Profit} = \{(P - X) - E(L)\}/P \qquad (14.2)$$

A 5% profit measure seems eminently reasonable, at least at first first glance. But on second thoughts, it becomes curiouser and curiouser. If we draw on the analogy between insurance polícies and debt, then the insurer is "borrowing" from its policyholders. Normally, with borrowing, the borrower pays interest to the lender. But with the 5% margin, the insurance company is not only borrowing its policyholders money but is graciously <u>accepting</u> 5% interest. The premiums are then invested and the insurer makes a second level of income on this debt. Nice work if you can get it. Suppose that the investment income was 10%. Does this mean that the insurance company makes a return of 15% (5% + 10%) for its investors. Not at all. The riddle gets deeper.

Investors measure their rate of return, not on the premium income, but on the amount they invest; i.e. the surplus. Since the insurer usually writes perhaps two, or even three, dollars of premium for every dollar of surplus, a profit of 15 cents per dollar of premiums might translate into two or three times that amount when related to surplus. Thus what seems like a modest rate of underwriting profit could well be a fabulous rate of return on equity. But still the story is not finished. Suppose the line is medical malpractice or

some other line in which claims typically take several years to settle. The insurer does not simply invest the premiums (net of expenses) for one year at 10% but for several years. Thus, the return on equity is multiplied upwards yet again. Quite possibly, an innocuous 5% profit could be a license to print money!

Is this is being a little dramatic? Well let us try an illustration.

*********************************ILLUSTRATION*********************************

We will first use the example to show what expected return on equity is implied by the 5% underwriting profit ratio. The Syzygy Insurance Company is being kept in line by Stata regulators who wish to impose a price structure yielding a 5% underwriting profit. In this example, we assume that there is the probability of ruin is zero. This simplifies the calculation considerably (But if the assumption were relaxed, the rate of return on equity that is estimated will be understated) example uses the following numbers;

| | |
|---|---|
| Premiums P | \$100m |
| Expenses X | \$ 20m |
| Expected losses E(L) | \$ 75m |
| Expected return on investment E(r) | 0.15 |
| Surplus S (40% of premiums) | \$ 40m |
| Turnover cycle n | 2 years |

Note first that the equity or surplus has been expressed as a ratio of premiums implying that this value is determined as the leverage decision. Notice also that these values, when substituted into equation 14.2, yield an expected underwriting profit of 5% which is the traditional "reasonable" level, i.e.

$$\text{Underwriting profit} = (100 - 20 - 75)/100 = 0.05$$

Now, in calculating the expected rate of return on equity, bear in mind that the insurer is able to invest the opening cash flows (net of expenses) and will earn investment income. Since the turnover cycle is 2 years, the funds will be invested for two years before they need to be liquidated to pay losses. Using equation 13.7,

$$[1 + E(r_e)]^n - 1 = \frac{(S + P - X)(1 + E[r])^n - L - S}{S}$$

$$[1 + E(r_e)]^2 = \frac{(40 + 100 - 20)(1 + 0.15)^2 - 80 - 40}{40}$$

$$= 0.9675$$

Thus,

$$E(r_e) = 0.4027$$

The 5% expected underwriting profit, in this case, yields a 40% expected rate of return on equity!

****************************************************************************

While this is a simple example, the parameters given are in no way outrageous. The issue would seem to be that this level of profit could not be sustained in a competitive market. With such an expected rate of return on equity investors would flock to this industry starting up new firms; while existing firms would try expand their output. Such intense competition would drive down the price and, with it, the expected rate of return on equity. This fall in prices will continue until they just offer an adequate return to equityholders. If, on the other hand, prices are too low, capital will leave the industry in search of higher returns elsewhere. Supply will thereby contract and the resulting excess demand will drive up prices until the expected return on equity matches that required by investors.

Some sixty years after the National Association of Insurance Commissioners (NAIC) recommended the 5% standard, an advisory body to the NAIC recognised that that it was without theoretical foundation. Their report also recognised that, to derive the competitive level of premiums, it is necessary to look to the expected return on equity.

## SETTING COMPETITIVE PREMIUM RATES

The cash flow model developed in the previous chapter now can be used to set a competitive price for insurance policies. By definition, such a price will offer the investors who contribute the insurance company's surplus a competitive expected return on equity. So let us start with the rate of return on equity. This has been calculated in two ways as we have proceeded throught his chapter. If there is no default risk, the expected return on equity can be taken from equation 13.7

$$[1 + E(r_e)]^n - 1 = \frac{(S + P - X)\{1 + E(r)\}^n - E(L) - S}{S} \tag{13.7}$$

This formula worked without default risk because we did not have to worry about negative values for terminal equity T. But with default risk, it was necessary to calculate each possible value of T, substituting zero where T would otherwise have been negative. This approach was used earlier in the example of The Gossamer Insurance Company. Here we will assume zero default risk and use equation 13.7 to derive the competitive price. To solve for the competitive price it would be necessary to estimate the rate of return required by the owners of the firm to compensate them for the risk of holding the insurance firm's stock. Rearrangement of equation 13.7 is a little cumbersome, but a solution for P results as follows;

$$P = S[\{(1 + E[r_e])/(1 + E[r])\}^n - 1] + E[L]/\{1 + E[r]\}^n + X \tag{14.3}$$

**********************************ILLUSTRATION********************************

To see the effect of this formula, the previous example of the Syzygy Insurance company that was subject by regulation, to the 5% underwriting profit. The premiums of $100m charged by Syzygy for its portfolio implied an underwriting profit of 5% but an expected rate of return on equity of 40.27% What level of premiums would be necessary to secure a competitive return on equity and what would the implied underwriting profit be. Suppose that investors require an expected rate of return of 17% to invest in Syzygy. This could be derived from the Capital Asset Pricing Model by estimating the beta for Syzygy. i.e.

$$E(r_e) = r_f + \beta[E(r_m) - r_f]$$

From equation 14.3, the relevant premium can be estimated to deliver the 17% expected rate of return on equity.

$$\begin{aligned} P \quad = \quad & \$40m[\{(1 + 0.17)/(1 + 0.15)\}^2 \; - \; 1] \\ + \quad & \$75m/(1 + 0.15)^2 \quad - \quad \$20m \\ = \quad & \$78.11m \end{aligned}$$

You may check this answer by plugging P = $78.11m into equation 13.7 and confirming that the expected rate of return on equity is 17%.

Now consider what rate of underwriting profit is implied by this price.

$$\begin{aligned} \text{Underwriting Profit} &= \{(P - X) - E(L)\}/P \\ &= \{78.11 - 20 - 75\}/78.11 \\ &= -21.62\% \end{aligned}$$

******************************************************************************

The underwriting profit required to deliver a competitive rate of return to the owners was negative. Or put more succinctly, <u>the competitive rate of expected underwriting profit is negative</u>. This seems very confusing. Moreover, it certainly perterbs many in the insurance industry. One frequently encounters statements in the press or even in the annual reports of insurance firms in which attention is drawn to negative underwriting profits. The implication often draw is that the insurer is losing money by selling policies at such a price. This does not follow. The confusion is in the word profit. The underwriting profit is not profit at all. A better analogy is that it is the <u>interest</u> rate paid by insurers for their use of the policyholders' funds. It would be silly to say that a bank is losing money because it pays interest to its depositors. It is equally silly to imply that insurers lose money by "suffering" an underwriting loss.

Look again at equation 14.3.

$$P = [((1 + E[r_e])/(1 + E[r]))^n - 1] + E[L]/(1 + E[r])^n + X \qquad (14.3)$$

If it were not for the first term on the right hand side, this formula would appear to be identical to that used in Chapter 4 and requoted as equation 14.1 above. The second and third terms are simply the discounted values of future loss payments and expenses. But there is a difference here. The discount rate used in 14.3 for losses is the expected rate of return on the insurance company's investment portfolio. But in chapter 4, we suggested using a risk adjusted discount rate which was appropriate to the particular risk in the loss payments.

There is, in fact, no conflict. The rather awkward first term in equation 14.3 is a correction for the fact that we are using a different discount rate. This can be seen most clearly by imagining that the insurance firm was able to successfully diversify all insurance risk by the law of large numbers and invested in a riskless investment portfolio. Such a stable beast would be a riskless insurance company. Investors would require only the risk free rate to invest their captital in such a firm. Thus $E(r_e) = E(r) = r_f$. The first term in equation 14.3 neatly cancels out to zero and we are left with;

$$P = E[L]/(1 + r_f)^n + X \qquad (14.4)$$

This is identical to the approach in chapter 4. The discount rate for expected losses is the risk free rate which is quite appropriate since aggregate losses are now risk free.

There is another way of calculating competitve, or fair, insurance premiums; using the capital asset pricing model (CAPM). This alternative is illustrated in the Appendix to this chapter. In particular we show there that both the cash flow method and the CAPM are compatible and can produce the same results.

## INSURANCE PRICES AND INFLATION

### (a). "Past" and "Future" Inflation

We will probe deeper into the pricing of insurance policies. In particular, we will examine how inflation affects insurance prices. Recall that insurance contracts give rise to cash flows at different points in time. Some of these cash flows are affected by changes in price levels price levels after the contract is written. Thus the policyholder and insurer must anticipate the effects of FUTURE inflation on the adequacy of their insurance contract. We will first clarify the timing of the cash flows under insurance contracts and show the affect of changing price levels. In particular we will distinguish past and future inflation. We will then show how future price changes affect the profitability of the insurance firm and how such changes may be anticipated when setting premims.

An insurance policy provides cover for events that arise in over an defined period. If a loss arises, a payment is made to the policyholder. Sometimes, the payment is an agreed sum which is specified in the contract. Life insurance usually has such an agreed sum. The sum usually will be selected so that the policyholder receives an adequate payment for the loss. If price levels change between the time the policy is written and the time the loss is settled, the value of the payment will be eroded. In other policies, the insurance promise is to pay a sum which is necessary to compensate the policyholder for the financial loss suffered. Property and liability insurance usually works on the compensation or indemnity principle. Inflation is important in these policies since the insurer must anticipate future price levels when setting premiums.

FIGURE 14.1

```
  Period Preceding    Contract Year     Run Off of Claims
   Contract Year
<----------------><-------------><----------------->          ----->time
<  PAST INFLATION  ><            FUTURE INFLATION         >

0                  1                      2
```

TIME 2

An insurance claim is settled
(Note that the loss will arise in the contract period but the final settlement of the claim will take place subsequently)
The adequacy of the payment will depend on price levels at time 2

TIME 1

Insurance contracts are written
Insurers will set premiums
Policyholders will determine the adequacy of the value insured
At this time, policyholders and the insurer will know the CURRENT level of prices. But in determining the adequacy of premiums and the adequacy of the value insured, they must anticipate FUTURE price levels prevailing when the loss is settled, i.e. time 2

TIME 0

The date on which the previous year's contracts were written.
Premiums and values insured would have been set knowing time 0 prices but anticipating future price levels.

Over the whole time period shown in Figure 14.1, prices may change. We will separate the period before the contract is written from the period after the contract is written. PAST INFLATION describes price changes before time 1. Past inflation is not directly relevant in determining premiums to be charged at time 1. FUTURE INFLATION describes price changes after the contract is written, i.e. after time 1. The profitability of the insurer may be affected by future inflation and expected future inflation is important in setting premiums at time 1. The concern here is with future inflation. But we

will see that its effects differ according to whether the policy pays an agreed nominal value or whether it offers compensation in real values.

(b). Policies Written in Nominal Values and Policies Written in Real Values

With some insurance contracts the nominal amount to be paid in the event of a loss is agreed in the contract. Many life insurance policies are agreed value contracts, though some permit the policyholder to share in the investment performance of the insurer in various ways. This investment sharing is usually an "add on" to a basic agreed value life insurance policy. Some policies for valuables are also agreed value policies. For example, the value of an expensive ring or necklace might be agreed at inception, and that value will be paid in the event of loss. With such policies, future inflation does not affect the amount to be paid.

With other policies, the policy agrees to indemnify the policyholder should a future loss arise. The amount required to provide full compensation will depend upon the value or replacement cost etc. at the time the loss is t be paid. For example, a policy might cover physical damage to an automobile for one year. If a loss occurs towards the end of the policy year, then the cost to settle the loss will depend upon repair costs at that future time. Settlements are affected by inflation during the life of the contract. The insurer should therefore try to anticipate future inflation when writing the contract. The insurer is even more vunerable on liability insurance. Many liability claims take years to settle. Thus, when writing insurance contracts and setting premiums, the insurer may have to anticipate future inflation in court awards for several years from the inception date.

(c). Future Inflation and Real Valued Policies

We will introduce the issues by way of example. Consider first a world without inflation. For each dollar of premiums, the insurer holds 50 cents surplus. Losses are expected to be \$1.05 per dollar of premiums. For simplicity, expenses will be ignored, (if you prefer, the premiums may be assumed to be specified net of expenses).The insurer invests its reserves at a rate of return of 5% and the turnover cycle is one year. The expected underwriting profit is;

$$\{P - E(L)\}/P = \{\$1 - \$1.05\}/\$1 = -0.05$$

The terminal value of equity is;

$$E(T) = (\$0.5 + \$1)(1 + 0.05) - \$1.05 = \$0.525$$

which gives an expected rate of return on equity of;

$$E(r_e) = (\$0.525 - \$0.5)/\$0.5 = 0.05$$

We now introduce the effects of inflation. Let us define $E(L_r)$ to be the expected value of losses if these losses were to be settled at the price levels prevailing at the time the contract is written. In the above example,

the price was implicitly based on the assumption of zero inflation, i.e. on the expected real value of losses $E[L_r] = \$1.05$. Now, suppose that, in spite of the <u>expection</u> of zero inflation, it turns out that prices actually rise by 10% between the writing of the contract and the time of loss settlement. With the 10% inflation rate, the actual nominal value of claim payments is $\$1.05(1+0.1) = \$1.155$ per dollar of premiums. What does this 10% inflationary increase in losses do for the underwriting profit and the return on equity. The actual calculated values are;

$$(P - L)/P = (\$1 - \$1.155)/\$1 = -0.155$$

$$T = (\$0.5 + \$1)(1 + 0.5) - \$1.155$$

$$= \$0.42$$

$$re = -16\%$$

Inflation, it appears, takes its toll on the insurance firm!

However, we can show that there is some inbuild protection from the apparent ravages of inflation. Consider that everybody EXPECTED the rate of inflation to be 10%. In such circumstances, investors would be unwilling to lend money at the 5% expected return assumed above. The expected real rate of return would be negative in these circumstances. Rather, investors would require an expected rate of return which compensated for expected inflation and still offered an adequate and positive real return on their investments. The relationship postulated by the economist Irving Fisher was that the expected nominal rate of interest $E(r_n)$ would impound both the expected inflation rate, $E[I]$, and the required real rate of interest $E(r_r)$ in the following fashion;

$$1 + E[r_n] = (1 + E(r_r))(1 + E[I]) \qquad (14.5)$$

In Fisher's relationship, the real rate of interest is assumed to be constant over time and we shall simply refer to this as $r_r$ If this relationship holds, the expected nominal interest rate in our example would be;

$$E(r_n) = (1 + 0.05)(1 + 0.1) - 1 = 0.155$$

With insurers investing their reserve funds at the expected nominal rate of interest of 15.5%, instead of the previous 5%, the expected terminal value of equity and the expected rate of return on equity are now recalculated as follows;

$$E(T) = (\$0.5 + \$1.0)(1 + 0.155) - \$1.155$$

$$= \$0.5775$$

and

$$E(r_e) = (\$0.5775 - \$0.5)/\$0.5$$
$$= 15.5\%$$

Do not be tempted to think that the insurer is better off because of inflation. The "no inflation" expected return on equity (i.e. the expected real return) was 5%. With expected inflation of 10%, the expected return on equity is 15.5%. But these are NOMINAL values. The REAL expected rate of return on equity is still 5% as shown by the Fisher formula;

$$1 + 0.155 = (1 + r_r)(1 + 0.1)$$

which gives;

$$r_r = 0.05$$

Which is exactly the same real expected return as in the no inflation case! The expected return on equity was unchanged by the introduction of expected inflation.

We will now formalize these ideas to derive a set of propositions concerning the effects of inflation on insurance operations. The simple pricing formula established in Chapter 4 was

$$\text{Price} = \text{NPV} = \Sigma\, C_i/(1+k)^i \qquad (14.6)$$

where $C_i$ = expected cash flow at time i
$k$ = risk adjusted discount rate

We saw in the last Chapter that this price implied a competitive expected return on equity. Think of both the cash flows and the discount rate as specified in nominal terms. The discount factor can be broken down into real and inflation components as shown in the Fisher equation, 14.5.

$$k = E(r_n) = (1+r_r)(1+E(I)) \qquad (14.7)$$

The expected cash flows also may be expressed to reveal the effects of inflation. For example, a loss expected in one year is;

$$C_1 = E(L_r)(1+E(I))$$

Thus, on a one year turnover cycle, the competitive premium is calculated as;

$$P = X + \frac{E(L_r)(1+E(I))}{(1+\bar{r}_r)(1+E(I))} \qquad (14.8)$$

$$= X + \frac{E(L_r)}{(1+\bar{r}_r)}$$

The expected inflation items simply cancel out. The competitive premium is the expected real value of losses discounted at the real discount rate. This is a very powerful result. It is also an apparently surprising result but it

does yield to commonsense. It tells us that insurance firms do not have to change their premiums when they think that inflation rates will change. It is not necessary to raise premiums when price levels are expected to rise. True, the nominal value of expected claims will rise. But inflation expectations are also built into expected returns on investment and this changes the discount factor. The two effects cancel out. The insurer is worse off by virtue of the effect of expected inflation on losses but better off because of the effects of expected inflation on premiums. Insurers are hedged against the effects of expected inflation.

There are two important qualifications that must be made to this analysis. The first concerns differences between the general rate of inflation for the economy as a whole and the specific price indices that affect insurance firms. The second concerns the difference between expected and actual inflation.

A price index shows the average changes in prices for a bundle of goods; thus the retail price index shows the average change in prices for a bundle of consumer goods whereas the wholesale price does the same thing for producer goods. To the extent that interest rates embody inflation expectations, we would expected these rates to reflect the combined effect of all price changes in the economy. However, insurance loss payment are affected by much more specific price changes, such changes in the index of liability awards, hospital costs or automobile repair costs. There is no reason why the specific price changes should be identical to the composite index of which they are but a small part. Thus the inflation rates for losses and the inflation rate emdodied in the investment expected return may be quite different.

$$P = \frac{E(L)(1+E(I_L))}{(1+r_r)(1+E(I))} \qquad (14.9)$$

where $I_L$ is the inflation rate for losses rate

Now, inflation will affect premiums insofar as the general rate differs from the specific loss rate. If the latter is expected to be higher, then premiums will rise. But, if the general rate is higher than the loss rate expected, premiums will fall in spite of the inflationary conditions. Recent experience in most property-liability lines of insurance shows that the loss inflation rate has exceeded the general rise in consumer prices

**********************************INSERT******LOSS INFLATION >RATES****

The second qualification to our analysis concerned actual, as opposed to expected, inflation. Contracts are priced when they are undertaken even thought the insurer does not know how much must eventually be paid in claims. While expected inflation may be built into the price in accordance with equation 14.9, the insurer may gain or lose if actual inflation turn out to be different from that expected. If insurance contracts were priced on the expectation of 10% inflation but actual inflation turns out to be 15%, then the insurer is the loser. The unexpected component of inflation turns out to be a windfall loss to the insurer. But the reverse also could occur. Lower

than expected inflation will bring a windfall gain

******************************ILLUSTRATION**************************************

The Equilibrium Insurance Company is pricing its portfolio of property policies. The expected losses, if settled at today's prices, are 100, expenses are 20 and the firm has 50 in equity. Based on past experience, losses have been increasing at a faster rate than the consumer price index. The expected inflation rate for losses is therefore estimated at 10% compared with an expected increase of 7% in the consumer price index. The nominal interest rate is 12%. Equilibrium sets its prices as follows (equation 14.9),

$$P = 20 + \frac{(100)(1+0.1)}{(1+0.12)}$$

$$= 118.21$$

Now consider the expected return on equity, assuming that the insurer also invests at the nominal rate of 12%.

$$E(T) = (50 + 118.21 - 20)(1+0.12) - 100(1+0.1)$$

$$= 56$$

and

$$E(r_e) = (56 - 50)/50$$

$$= 12\%$$

But what if the actual inflation rate for losses turned out to be 5% in spite of the expected 10%? The insurer would make a windfall profit. The ACTUAL return on equity would be;

$$T = (50 + 118.21 - 20)(1+0.12) - 100(1+0.05)$$

$$= 61$$

and

$$r_e = (61 - 50)/50$$

$$= 22\%$$

The insurer has made a windfall gain. But things could easily have gone in the opposite direction. If the inflation rate for losses turned out to be 15%, you may calculate that the return on equity would have been 1.99%.

******************************************************************************

### (d). Future Inflation and Policies Written in Nominal Values

We have seen that, with policies that cover real values, the insurer has an inbuilt hedge against expected future inflation. Inflation will simultaneously increase the expected loss to be paid in the future and the

discount rate. The effects offset each other. But it is not a perfect hedge since the general inflation expectation impounded in interest rates may differ from the specific inflation expected for losses. Moreover, the insurer can make windfall gains or losses if actual inflation turns out to be less than (more than) expected inflation. We now consider policies in which the amount to be paid is fixed in nominal value such as a life insurance policy.

Consider a very simple life insurance policy which offers one year of term life cover for a fixed premium. Under such a policy, and indeed many other life policies, the face value is specified in nominal terms. If the policyholder arranges a policy with a face value of $100,000, that is paid if he dies. The competitive premium for such a policy is:

$$P = X + \frac{(\text{probability of death}) \times (\text{face value})}{(1 + r_r)(1 + E(I))} \tag{14.10}$$

Inflation only enters the premium calculation at one point; in the discount rate. Inflation, after the contract has been written, has no affect on the value to be paid since this already has been agreed. Thus when inflation expectations are high, leading to high nominal interest rates, insurers will tend to reduce their premiums for policies that are denominated in nominal values. Conversely, a fall in inflation expectations will tend to reduce interest rates, thereby reducing the discount rate in equation 14.10 and increasing the competitive level of premium. Nominal valued policies, unlike real value policies, are not hedged against future inflation and the premiums for such policies will tend bear an inverse relationship to current interest rates

************************************ILLUSTRATION**********************************

The Disequilibrium Life Insurance Company is revising its premium schedules for term life cover. A one year term with a unit face value of $1000 is the focus. Last year, premiums were set on the basis of the then current interest rate of 8%. Thus for an individual with a mortality probability of 0.002 and 20 cents expenses,

*needspace(4 lines)

$$P = \$0.2 + \frac{(0.002) \times (\$1000)}{(1 + 0.08)}$$

$$= \$2.05$$

Now, interest rates have risen to 12%, reflecting an upward revision in the inflationary expectations held by investors. The new premium will be:

$$P = \$0.2 + \frac{(0.002) \times (\$1000)}{(1 + 0.12)}$$

$$= \$1.99$$

*********************************************************************************

This example implies that life insurance firms would be constantly changing their premiums as changes in inflation expectations affected discount rates. In fact, this probably has not been the case, but it does allow some insights into recent changes in life insurance markets. It is widely believed that insurance company actuaries tend to use very low discount rates when calculating life insurance premiums.[2] Such conservative discounting implies that life insurance policies do not offer policyholders an attractive investment return and life insurance would not fare well alongside alternative investment vehicles. This disadvantage to life insurance would be particularly severe at times when individuals could invest in capital markets at very high expected rates of return. Such a period was the early 1980's. The response of the industry was to design new policies, such as universal life insurance. These policies include some insurance protection but they also offer the policyholder a variable investment benefit that depends on the investment performance. In this way, variations in capital market rates accrue to policyholders, not so much in changes in the basic premium for pure life insurance, but in terms of the value of the investment vehicle that is attached to the basic life insurance.

## PROFIT CYCLES IN PROPERTY LIABILITY INSURANCE

An interesting feature of insurance markets that has puzzled observers for many years is the "underwriting cycle". Underwriting profits are not stable. But the instability is not a purely random movement. Rather, underwriting profits <u>appear</u> to move in a cyclical pattern, with a period of about six years to complete each cycle. Much has been written on the cycle in the professional literature, though the cycle has only recently been subject to serious economic analysis.[3]. The cycle is still not well understood, but we shall see that the previous analysis of insurance pricing does give some definite insights.

First, the cycle needs to be defined. The cycle appears to be a non random movement in underwriting profits over time. Underwriting profits appear to oscillate in a six year cycle. If such a cycle does exist, it is clearly not a perfect and regular oscillation. But then again, the movement does not appear to be random. Figure 14.2 shows underwriting profits for U.S. property liability insurance for the extended period 1955 - 1985. Various writers have tested this series using various statistical techniques and have concluded that the movement does resemble a six year cycle.[4]. If this movement really is a cycle then why does such a cycle exist?

First, consider what it is that is being measured; the underwriting profit. This was defined in the previous chapter, equation 13.5;

$$\text{Underwriting Profit} = \{(P - X) - L\}/P$$

It should be clear now that this is not a true measure of the profit of the insurance firm since it ignores any investment income. But it is also important to note that insurers usually record underwriting profit using undiscounted values. Premiums are premiums earned during the accounting year and expenses are those paid duing the year. In each case, no adjustment is made for whether the cash flows arise at the beginning or end of the year. Similarly with losses. Losses incurred during the accounting year may be paid

during that year in which case the payments are recorded. But it the loss is unsettled at the end of the year and estimate is made as to its ultimate expected cost (usually in nominal values) and this estimate is added to the paid losses. Even though outstanding losses may be paid months or even years into the future, the outstanding claims estimate usually is an undiscounted value.

Now consider that insurance firms are correctly pricing their policies in order to offer a competitive expected return on equity, i.e. they price by equation 14.9. For simplicity, we assume that the expected inflation rate for losses is the same as the general inflation rate. What is the EXPECTED level of underwriting profit?

$$\text{Expected Underwriting Profit} = \{(P - X) - E(L)(1 + E(I))\}/P$$

$$= 1 - \frac{\{X + E(L)(1 + E(I))\}}{P}$$

but, since P is defined b is defined by equation 14.4.

$$\text{Expected Underwriting Profit} = 1 - \frac{\{X + E(L)(1 + E(I))\}}{E(L)/(1 + r_r)}$$

$$= -r_n - \frac{X}{E(L)(1 + r_r)} \qquad (14.11)$$

where $r_n$ is the nominal interest rate and $r_r$ is the real interest rate. Now, according to the Fisher theory, the real rate of interest is constant but changes in inflation expectations will cause sympathetic changes in nominal interest rates. Thus expected underwriting profits would change as nominal interest rates change as shown in equation 14.11. <u>Underwriting profit will be expected to fall as nominal interest rates rise and vice versa.</u> This is an interesting and important result for the following reasons.

1. Equation 14.11 leads to the prediction that there will be an inverse relationship between expected underwriting profits and nominal interest rates. This relationship cannot be tested directly since <u>expected</u> underwriting profits cannot be observed. We can only observe <u>actual</u> underwriting profits recorded at the end of each accounting period. However, Figure 14.2 also shows the nominal interest rate at the beginning of each accounting year. It does indeed seem that there is a tendency to an inverse relationship. More sophisticated statistical analysis supports this relationship.[5]

2. The inverse relationship between nominal interest rates and underwriting profit was established on the assumption that insurance policies were priced to give a competitive rate of return on equity. Recall that equation 14.9 was used to derive equation 14.11. This implies that <u>underwriting profits must move inversely to interest rates in order to maintain a competitive rate of return on equity.</u>

There appears to be a strong view in the industry that underwriting profits should be stable over time. The same view holds that changes in target underwriting profits in response to changing interest rates, is unsound business practice. Such practice is termed "<u>cash flow underwriting</u>", a term that to many minds is almost synonymous with imprudent underwriting. The analysis given here directly contradicts that view.

3. Changes in the level of underwriting profit do not necessarily imply that premium levels have changed, or should change. If inflation of losses is the same as general inflation rate, the competitive level of premium is unaffected by changes in nominal interest rates. This result was shown earlier in equation 14.8. Thus while an increase in nominal interest rates would lead to a decline in competitive rates of underwriting profit,(equation 14.8), this would not affect the competitive levels of premium (equation 14.11). Two qualifications must be made. First, differences between the expected general inflation rate and the expected loss inflation rate may require some adjustment to premiums. This was shown in equation 14.9. Second, if the policy is denominated in nominal rather than in real values then changes in inflation expectations do affect the competitive level of premium. This was shown in equation 14.10.

Where does this leave the discussion of the cycle? What has been explained is that underwriting profits would not be constant over time in a competitive market. Underwriting profits would be expected to change as inflationary expectations changed which thereby changed interest rates. This explains <u>movement</u> in the rate of underwriting profit over time but it does not explain why the movement appears to follow a six year cycle. For the apparent cyclicality of the movement, we have no satisfactory explanation. This may refect the inadequacy of current research. Or it may reflect that the cycle itself is an illusion. Attempts to explain the cycle may well have caused researchers to be over-enthusiastic in their search for patterns of movement. In much the same way the early astronomers (notably Percival Lowell) searched for patterns in the features of Mars. That search led to the idea of canals on Mars. The moral is that if we search hard enough for patterns we will find them. But it is difficult to know whether the exist more in our imagination than in reality.

Our second conclusion about the cycle is that it may be "much ado about nothing". Many conclusions have been drawn about the changing economic health of insurance firms just by looking at changing underwriting profits. Such conclusions are often used to justify price changes. Analysis conducted at this level is trivial and misleading. Underwriting profits alone tell nothing of the profitability of insurance firms and carry no implications for price changes. The fact that underwriting profits do respond to changes in the capital markets should be cause for reassurance not alarm. It shows that the insurance industry's judgements about future inflation are not entirely out of line with those of investors and it shows that the insurance industry is properly reflecting the time value of money in its premium calculations. Anyone who disagrees with this conclusion might examine how he would feel if banks failed to change their interest rates to depositors when the general level of interest rates rose.

----------

1. If it were true that future inflation could be predicted, e.g. be extrapolation, from past inflation, then past inflation might be indirectly relevant in setting premiums.

2. Some indirect evidence of this could be seen until recently in the very low rates that policyholders could borrow from the insurance company against the surrender value of life insurance policies.

3. see for example, Venezian, [1985], Cummins and Outreville, [1987] and Doherty and Kang, [1986]

4. Smith [1981], and Doherty and Kang [1986] use spectral analysis to show that the movement of underwriting profits is compatible with a six year cycle. This result is supported by Venezian [1985] and by Cummins and Outreville [1987] using quadratic regression analysis

5. see Doherty and Kang [1986]

## REFERENCES

Cummins, J. D. and D. J. Harrington, (eds.),(1986), "Fair Rate of Return in Property-Liability Insurance", Kluwer-Nijhoff Publishing, Norwell, Mass..

Cummins, J. D. and D. J. Nye, (1981), "Portfolio Optimization Models for Property Liability Insurance Companies", Management Science, 27, 414-430

____________________________, (1981), "Inflation and Property-Liability Insurance", in J. D. Long (ed), Issues in Insurance, Vol. 2. American Institute for Property and Liability Underwriters, Malvern, Pa..

Cummins, J. D. and F. Outreville

Doherty, N. A. and J. G. Garven, (1986), "Price Regulation in Property Liability Insurance", Journal of Finance, 41, ****************************

Doherty, N. A. and H. B. Kang, (1987), "Price Instability for a Financial Intermediary: Interest Rates and Insurance Price Cycles" Working Paper, Wharton School, University of Pennsylvania

Fairley, W. (1979), "Investment Income and Profit Margins in Property-Liability Insurance: Theory and Empirical Results"' Bell Journal of Economics, 10, 192-210

Hill, R. (1979), "Profit Regulation in Property-Liability Insurance", Bell Journal of Economics, 10, 172-191.

Kraus, A. and S. A. Ross, (1982), "The Determination of Fair Profits for the Property-Liability Insurance Firm", Journal of Finance, 37, 1015-1028.

Stewart, B. D.

Venezian, E.

## APPENDIX

### DERIVING THE COMPETITIVE PREMIUM WITH THE CAPITAL ASSET PRICING MODEL

The cash flow model for setting competitive insurance premiums was illustrated with an example of the Syzygy Insurance Company. In looking at the Syzygy example, we considered the possibility that the expected rate of return required by investors to contribute surplus to the insurance firm might be estimated by the capital asset pricing model (CAPM). i.e.

$$E(r_e) = r_f - \beta[E(r_m) - r_f]$$

If the CAPM does apply then we may use an alternative formula for calculating the competitive price. This alternative formula does not contradict the previous approach, equation 14.3. Rather it is simply a special

case of equation 14.3 which applies if assets are priced according to the CAPM.

When first looking at the CAPM in Chapter 9, the idea of value additivity and separable decisions was introduced. Suppose a firm has two internal operations. If those operations are separable, then the cash flows from each may be valued independently, as though the the units were themselves separate firms traded on the market. The insurance firm may be thought of as two internal operations, insurance operations and investment operations. The investment operations are, by definition, conducted on the capital market and it is reasonable to use the CAPM to value those activities and to determine the competitive rate of return. But with separation, we can use the same model, the CAPM, to establish the value of the insurance operation, i.e. to yield the competitive premium.

The simplest way to present the CAPM solution is to express the competitive rate of underwriting profit, which we will denote $E(r_u)$, We will then use this to derive the premium. The solution for the competitive rate of underwriting profit is;

$$E(r_u) = -(n)(r_f) + \beta_u[E(r_m) - r_f] \tag{14.A.1}$$

where $\beta_u$ is the insurance beta

which is calculated in the usual way, i.e. $= \mathrm{Cov}(L, r_m)/\sigma^2(m)$

This is the familiar CAPM equation but with a couple of subtle applications. The intercept is negative. The reason is simply to reflect the direction of the cash flow; the firm "borrows" from its policyholders rather than lends to them. The second change in the intercept is the multiplication by the turnover cycle "n". This simply reflects that the transaction may take several years. The CAPM without such adjustment applies only to one year. Once $E(r_u)$ is calculated, the premium is then solved from

$$E(r_u) = \{(P - X) - E(L)\}/(P - X) \tag{14.A.2}$$

This is a slightly different definition of underwriting profit from that used earlier. The difference lies in the treatment of expenses. The general definition of underwriting profit is the ratio of underwriting income to premiums. The CAPM approach shown in equation 14.A.1 actually defines underwriting profit as the ratio of underwriting income to premiums <u>net of expenses</u>. This is simply a quirk of the CAPM approach. If this equation is solved for the premium, the solution is

$$P = E(L)/\{1 - E(r_u)\} + X \tag{14.A.3}$$

**********************************ILLUSTRATION***********************************

We will estimate the competitive insurance price for Pterodactyl Aviation Underwriters Inc.. The expected losses from its current portfolio of \$60m. Underwriting, etc. expenses are \$10m and the turnover cycle is 2 years. The beta for the insurance portfolio is -0.2; the expected return on the market portfolio is 12% and the risk free rate 5%. The expected rate of underwriting profit from equation 14.A.1;

$$E(r_u) = -(2)(0.05) + (-0.2)[0.12 - 0.05]$$

$$= -0.114$$

The premium can now be solved from equation 14.A.3

$$P = 60/(1 - [-0.114]) + 10$$

$$= \$63.85m$$

*****************************************************************************

## A Comparison of the Cash Flow Pricing Model and the CAPM

We have just presented two models, each of which can be used to derive the fair or competitive premium rates for insurance policies. Which is the correct model? In fact, the models do not compete at all. They are not really different models at all. They are simply different presentations of the same ideas. Look again at the two models as presented in equation 14.3 and 14.A.3:

CASH FLOW MODEL PREMIUMS

$$P = E(L)/(1 + E(r)) + X + [((1 + E[r_e])/(1 + E[r]))^n - 1] \tag{14.3}$$

CAPM PREMIUMS

$$P = E(L)/(1 - E(r_u)) + X \tag{14.A.3}$$

The CAPM simply discounts expected losses at a discount rate, $E(r_u)$, which is determined by the risk free rate and the beta of losses. The discount rate looks a little strange since it is <u>deducted</u> in equation 14.A.3). But do not forget that the value of $E(r_u)$ as calculated from equation 14.4, usually is negative. Thus the double negative results in a positive rate. In contrast, equation 14.3 discounts expected losses at the discount rate "r" which is the expected rate of return on the insurance firm's investment portfolio. Of course, this discount rate may differ from $-E(r_u)$ which is used in equation 14.A.3. To compensate for this difference, the last term in equation 14.3 is added. Both approaches should therefore give approximately the same answer. To show this we will return to the above example of Pterodactyl Aviation Underwiters. This was calculated with CAPM. we will now recalculate with the cash flow method described by equation 14.3.

Equation 14.3 requires more information than we have here. We need the expected return on investment, E(r), and the expected return required by investors to hold the equity of the insurer $E(r_e)$. To estimate each of these requires that we know the respective betas. Suppose that the investment beta, $\beta_i$, is 1.0. From this we may estimate two things. First, the expected rate of return on the investment portfolio will be

$$E(r) = 0.05 + 0.6[0.12 - 0.05]$$

$$= 0.092$$

The second thing we may use the investment beta for, is to help derive the beta for the insurer's own equity, $\beta_e$. The insurance firm's shareholders will be exposed to systematic risk from one of two sources; the investment activities of the insurer and/or the underwriting activities. In fact, the equity beta is the weighted average of the investment and underwriting betas;

$$\beta_e = [1 + (nP)/S]\beta_i + [P/S]\beta_u$$

The surplus is assumed to half the premium to maintain a level of leverage that is considered to be sound; i.e. S = 0.5P.

$$\beta_e = [1 + (2P)/(0.5P](0.6) \quad + \quad [P/(0.5P](-0.2)$$

$$= 2.6$$

Thus

$$E(r_e) = r_f + \beta_e[E(r_m) - r_f]$$

$$= 0.05 + 2.6[0.12 - 0.05]$$

$$= 0.232$$

Now we have all we need to solve equation 14.3;

$$P = S[\{(1 + E[r_e])/(1 + E[r])\}^n - 1] + E[L]/(1 + E[r]\}^n \quad + \quad X$$

$$P = (0.5)P[\{(1 + 0.232)/(1 + 0.092)\}^2 - 1] + 60/\{1 + 0.092\}^2 \quad + \quad 10$$

$$= 0.1364234P \quad + \quad 50.32 \quad + \quad 10$$

Thus

$$P = \$63.66m$$

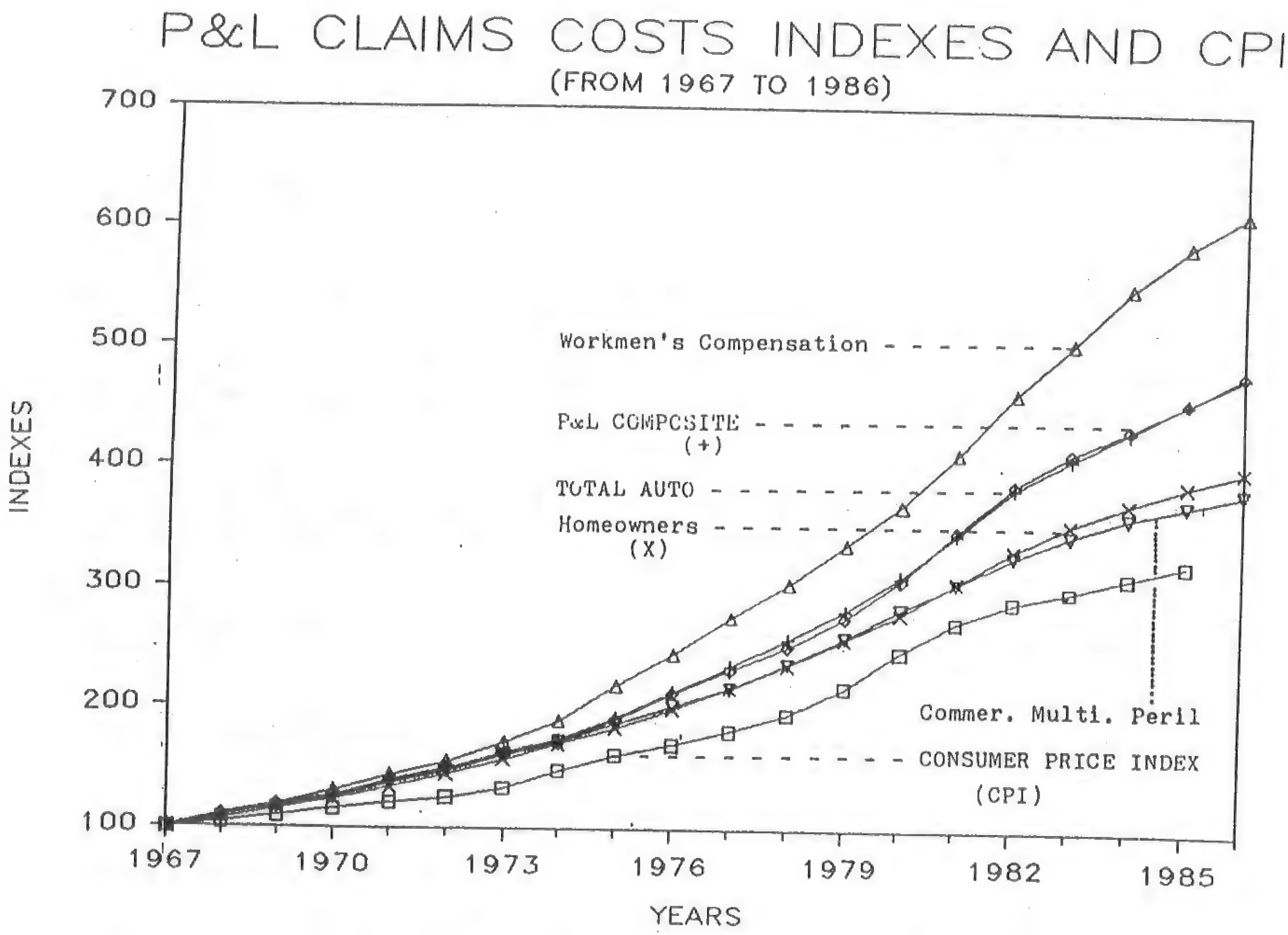
This differs slightly from the answer of $63.85 for Pterodactyl Aviation Underwriters. In the analysis just completed we had to add more information to solve equation 14.3. The additional information related to the level of surplus which was set at fifty percent of the total premiums. This was a somewhat rough estimate. A slightly different level of surplus would result in a more accurate answer.

UNDERWTITING RESULTS AND INTEREST(T-BILL) RATES
(FROM 1955 TO 1986)

| YEAR | *UNDERWRITING GAIN OR LOSS ($000) (1) | PREMIUM EARNED ($000) (2) | U.W.PROFIT (% TO PREM EARNED) [(1)/(2)]*100 | INTEREST (T-BILL) RATES (%) |
|---|---|---|---|---|
| 1955 | 590,973 | 10,033,006 | 5.890 | 1.753 |
| 1956 | 90,432 | 10,661,679 | 0.848 | 2.658 |
| 1957 | (130,333) | 11,525,936 | -1.131 | 3.267 |
| 1958 | 205,981 | 12,314,985 | 1.673 | 1.839 |
| 1959 | 426,979 | 13,379,120 | 3.191 | 3.405 |
| 1960 | 462,367 | 14,445,374 | 3.201 | 2.928 |
| 1961 | 484,089 | 15,142,579 | 3.197 | 2.378 |
| 1962 | 343,666 | 15,898,816 | 2.162 | 2.778 |
| 1963 | (165,557) | 16,446,749 | -1.007 | 3.157 |
| 1964 | (325,724) | 17,668,651 | -1.844 | 3.549 |
| 1965 | (352,264) | 19,077,274 | -1.847 | 3.954 |
| 1966 | 386,927 | 21,031,845 | 1.840 | 4.881 |
| 1967 | 193,619 | 22,816,851 | 0.849 | 4.321 |
| 1968 | (148,224) | 24,887,834 | -0.596 | 5.339 |
| 1969 | (529,899) | 27,711,016 | -1.912 | 6.667 |
| 1970 | 77,776 | 31,169,052 | 0.250 | 6.458 |
| 1971 | 1,381,801 | 33,920,246 | 4.074 | 4.348 |
| 1972 | 1,797,933 | 37,602,057 | 4.781 | 4.071 |
| 1973 | 791,735 | 40,892,726 | 1.936 | 7.041 |
| 1974 | (1,881,551) | 43,746,161 | -4.301 | 7.886 |
| 1975 | (3,593,960) | 47,844,414 | -7.512 | 5.838 |
| 1976 | (1,558,551) | 57,268,286 | -2.721 | 4.989 |
| 1977 | 1,926,315 | 68,870,281 | 2.797 | 5.265 |
| 1978 | 2,548,143 | 78,746,967 | 3.236 | 7.221 |
| 1979 | 23,606 | 86,959,718 | 0.027 | 10.041 |
| 1980 | (1,712,423) | 93,850,138 | -1.825 | 11.506 |
| 1981 | (4,463,887) | 97,554,652 | -4.576 | 14.029 |
| 1982 | (8,303,039) | 102,069,503 | -8.135 | 10.686 |
| 1983 | (11,087,953) | 107,195,857 | -10.344 | 8.52 |
| 1984 | (19,378,885) | 115,009,836 | -16.850 | 9.57 |
| 1985 | (22,503,429) | 133,341,852 | -16.876 | 7.47 |

DATA SOURCES:(1)BEST'S AGGREGATES & AVERAGES [P&L](ANNUAL)
(2)FEDERAL RESERVE BULLETIN (MONTHLY)

*UNDERWRITING GAIN OR LOSS BEFORE DIVIDEND TO POLICYHOLDERS


P&L CLAIMS COSTS INDEXES AND CPI
(FROM 1967 TO 1986)
INDEXES
700
600
500
400
300
200
100
1967
1970
1973
1976
1979
1982
1985
YEARS
Workmen's Compensation
P&L COMPOSITE
(+)
TOTAL AUTO
Homeowners
(X)
Commer. Multi. Peril
CONSUMER PRICE INDEX
(CPI)

U.S. CLAIMS COSTS INDEXES VS.CONSUMER PRICE INDEX(CPI)
(FROM 1967 TO 1986)
(1967=100)

| YEAR | TOTAL AUTO | AUTO BODI. | AUTO PROP. | AUTO PHYS. | W.C. | OTHER BODI. | OTHER PROP. | GLASS |
|---|---|---|---|---|---|---|---|---|
| 1967 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |
| 1968 | 110.8 | 108.5 | 112.4 | 113.5 | 109.6 | 108.2 | 107.9 | 105.6 |
| 1969 | 119.3 | 118.2 | 120.1 | 120.5 | 119.6 | 117.8 | 115.8 | 113.3 |
| 1970 | 126.5 | 128.3 | 125.5 | 124.6 | 130.9 | 128.3 | 122.5 | 120.5 |
| 1971 | 139.5 | 139.4 | 139.3 | 139.0 | 143.8 | 139.9 | 133.2 | 129.6 |
| 1972 | 149.6 | 147.1 | 150.9 | 152.2 | 155.0 | 147.7 | 142.1 | 135.2 |
| 1973 | 160.7 | 157.1 | 163.3 | 164.5 | 171.4 | 157.2 | 153.7 | 142.8 |
| 1974 | 171.1 | 171.6 | 172.7 | 172.3 | 189.0 | 171.8 | 161.9 | 149.5 |
| 1975 | 190.5 | 191.9 | 190.1 | 189.1 | 218.4 | 192.2 | 177.0 | 159.2 |
| 1976 | 212.1 | 214.5 | 210.3 | 210.2 | 244.7 | 214.8 | 197.2 | 175.6 |
| 1977 | 231.8 | 236.2 | 228.7 | 228.1 | 275.4 | 235.8 | 214.3 | 185.2 |
| 1978 | 251.8 | 258.5 | 249.6 | 244.3 | 304.1 | 256.9 | 232.7 | 202.2 |
| 1979 | 275.2 | 284.6 | 272.3 | 265.5 | 336.6 | 282.1 | 252.7 | 216.1 |
| 1980 | 306.9 | 316.3 | 304.3 | 297.8 | 368.5 | 314.8 | 271.3 | 234.1 |
| 1981 | 348.1 | 355.5 | 345.3 | 340.9 | 412.2 | 352.9 | 295.8 | 257.4 |
| 1982 | 386.3 | 400.9 | 381.2 | 371.4 | 462.8 | 399.3 | 322.1 | 275.6 |
| 1983 | 413.4 | 437.3 | 404.4 | 388.3 | 504.8 | 432.7 | 347.0 | 290.0 |
| 1984 | 434.2 | 472.6 | 419.4 | 394.3 | 551.9 | 457.8 | 367.4 | 299.7 |
| 1985 | 456.6 | 507.1 | 431.1 | 395.6 | 588.0 | 495.8 | 395.1 | 308.2 |
| 1986* | 480.7 | 538.3 | 451.8 | 410.9 | 615.4 | 535.8 | 406.5 | 320.9 |

| YEAR | THEFT | BOILER & MACH | FIRE | HOME OWNER | COMMER PERIL | INLAND MARINE | P&L TOTAL | CPI |
|---|---|---|---|---|---|---|---|---|
| 1967 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |
| 1968 | 106.9 | 105.2 | 106.3 | 107.0 | 106.4 | 106.6 | 109.6 | 104.2 |
| 1969 | 113.2 | 111.0 | 115.6 | 115.8 | 115.9 | 112.7 | 118.1 | 109.8 |
| 1970 | 119.0 | 116.8 | 123.8 | 123.4 | 125.0 | 118.4 | 126.2 | 116.3 |
| 1971 | 127.0 | 123.3 | 135.5 | 133.6 | 137.4 | 125.7 | 138.4 | 121.3 |
| 1972 | 132.9 | 131.9 | 146.0 | 144.0 | 147.0 | 132.1 | 148.5 | 125.3 |
| 1973 | 146.4 | 140.5 | 159.3 | 156.8 | 160.8 | 143.1 | 163.7 | 133.1 |
| 1974 | 155.5 | 153.9 | 172.4 | 170.4 | 172.6 | 153.4 | 173.8 | 147.7 |
| 1975 | 167.3 | 166.0 | 185.5 | 183.5 | 187.4 | 166.2 | 192.5 | 161.2 |
| 1976 | 180.3 | 181.7 | 199.8 | 199.9 | 202.0 | 179.1 | 213.4 | 170.5 |
| 1977 | 198.5 | 198.9 | 216.1 | 217.0 | 216.6 | 192.8 | 235.1 | 181.5 |
| 1978 | 215.9 | 218.2 | 235.1 | 235.9 | 236.9 | 209.2 | 257.5 | 195.4 |
| 1979 | 235.9 | 235.2 | 256.8 | 257.7 | 259.5 | 227.5 | 281.8 | 217.4 |
| 1980 | 257.1 | 256.7 | 278.3 | 278.7 | 282.6 | 250.0 | 310.8 | 246.8 |
| 1981 | 276.7 | 284.2 | 298.5 | 306.1 | 304.8 | 269.8 | 345.8 | 272.4 |
| 1982 | 295.0 | 295.2 | 317.1 | 332.8 | 327.4 | 287.1 | 383.1 | 289.1 |
| 1983 | 313.7 | 312.6 | 331.2 | 354.4 | 345.5 | 301.4 | 408.8 | 298.4 |
| 1984 | 338.2 | 336.5 | 342.1 | 371.9 | 360.5 | 319.7 | 432.3 | 311.1 |
| 1985 | 361.3 | 342.7 | 349.6 | 387.5 | 371.4 | 336.3 | 456.8 | 322.2 |
| 1986* | 318.6 | 351.9 | 356.3 | 400.5 | 381.7 | 352.2 | 479.2 | |

* ESTIMATED

DATA SOURCES: MORTON E. MASTERSON, Best's Review-P&L-(Sept.1976) and Best's Insurance Management Reports(Perspectives:5/25/'81 &4/14/'86).

© *The Journal of Risk and Insurance*, 1997, Vol. 64, No. 2, 231–270.



# Financial Risk Management by Insurers: An Analysis of the Process

Anthony M. Santomero
David F. Babbel

ABSTRACT

On-site visits to financial service firms were conducted to review and evaluate their risk management systems. In the insurance sector, this evaluation covered prominent life/health and property-liability insurers, both in the United States and abroad. The information obtained covered both the philosophy and the practice of financial risk management. This article outines the results of this investigation. It reports the state of risk management techniques in the industry. It reports the standard of practice and evaluates how and why it is conducted in the particular way chosen. In addition, critiques are offered where appropriate. We discuss the problems which the industry finds most difficult to address, shortcomings of the current methodology used to analyze risk, and the elements that are missing in the current procedures of risk managment.

## INTRODUCTION

The past decade has seen a dramatic rise in the number of insolvent insurers. The ostensible causes of these insolvencies were myriad. Some of the insolvencies were precipitated by rapidly rising or declining interest rates. Others resulted from losses on assets such as junk bonds, commercial mortgages, collateralized mortgage obligations, real estate, and derivatives. Mispricing of insurance policies, natural catastrophes, and changes in legal interpretations of liability and the limits of coverage hurt still others. The "churning" of policies by unscrupulous sales agents, insolvencies among the reinsurers backing the policies issued, noncompliance with insurance regulation, and malfeasance on the part of officers and directors of the insurance companies affected some as well. But despite the numerous and disparate apparent causes of these insolvencies, the underlying factor in all of them was the same: inadequate risk management practices. In response to this, insurers almost universally have embarked upon an upgrading of their financial risk

---

Anthony M. Santomero is Richard K. Mellon Professor of Finance and Director of the Financial Institutions Center at the Wharton School, University of Pennsylvania. David F. Babbel is Associate Professor of Insurance and Finance at the Wharton School, University of Pennsylvania.

The authors are grateful for the thoughtful comments and suggestions from the reviewer and Editor.

management and control systems to reduce their exposure to risk and better manage the amount they accept. In short, the industry has turned to financial risk management techniques as a way to improve performance.

Coincidental to this activity, and, in part, because of our recognition of the industry's vulnerability to financial risk, the Wharton Financial Institutions Center, with the support of the Sloan Foundation, has been involved in an analysis of financial risk management processes in the financial sector. During 1995 and 1996, on-site visits were conducted to review and evaluate the risk management systems and the process of risk evaluation that are in place.

In the insurance sector, system evaluation was conducted covering a number of prominent life/health and property-liability insurers, both in the United States and abroad. The information obtained on the philosophy and practice of financial risk management comes primarily through intensive interviews of these insurance firms, conducted by a team of researchers from the Wharton Financial Institutions Center.[1] Measured in terms of admitted assets, these firms range in size from $7 billion to well over $100 billion. They are organized as stock, reciprocal, or mutual insurers. Some firms restrict their activities to life insurance and pensions; others are multiline insurers, selling the full range of property-liability and life/health products. These visits were augmented by interviews conducted with additional large insurance firms domiciled in Japan and Europe as well as North America. As was the case above, these firms include life insurers, property-liability insurers, and a multiline company.

Our information was then supplemented by five published surveys. Three of these were the 1994, 1995, and 1996 "Insurer CIO Surveys," conducted by Goldman, Sachs & Co. (see Alexander et al., 1994; Siegel et al., 1995; and Millette et al., 1996). These surveys were based on responses from between 58 and 79 companies, depending on the survey, with approximately two-thirds from the life lines and one-third from the property-liability lines. The response rate in these surveys was roughly 40 percent. Over 90 percent of the responding companies had assets in excess of $1 billion. Another pair of surveys conducted by Lamm-Tennant (1995) and Lamm-Tennant and Gattis (1996) covered a much more extensive set of insurers, ranging in assets from $75 million to over $100 billion. There were 119 and 144 respondents to these 1995 and 1996 surveys, respectively, representing a response rate between 11 and 20 percent of the firms surveyed. The business mix of these firms was split about equally between life/health and property-liability insurance.

This article outlines the results of our investigation. It reports the state of risk management techniques in the industry—questions asked, questions answered, and questions left unaddressed by respondents. This report cannot recite the litany of approaches used within the industry, nor can it offer an evaluation of each and every approach. Rather, it reports the standard of practice and evaluates how and

[1] Members of the team included Anthony Santomero (leader), David Babbel, Yuval Bar-Or, Richard Herring, Paul Hoffman, Susan Kerr, Spencer Martin, Steve Pilloff, Jeffrey Trester, and Sri Zaheer.



why it is conducted in the particular way chosen. But even the best practice employed within the industry is not good enough in some areas. Accordingly, critiques will also be offered where appropriate. Our conclusion discusses the problems which the industry finds most difficult to address, shortcomings of the current methodology used to analyze risk, and the elements that are missing in the current procedures of risk management.

## Why Manage Risk—Some Generic Answers

It seems appropriate to begin our analysis of risk management techniques with a brief review of the reasons given for firm-level concern over the volatility of financial performance. The finance literature on why firms manage risk at all is usually traced back to 1984. In that year, Stulz (1984) first suggested a viable reason for objective function concavity, and his contribution is widely cited as the starting point of this burgeoning literature. Doherty (1985) provides the first comprehensive treatment of this topic for insurers in a finance framework. Since that time, several alternative theories and explanations have been offered. Santomero (1995) presents a useful review of these explanations. The reader is referred to that study for a thorough review and extensive citations to the economic literature.

The goal, as noted above, is to offer viable economic reasons for firm managers, who are presumed to be working on behalf of firm owners, to concern themselves with both expected profit and the distribution of firm returns around their expected value. The rationales for risk aversion can usefully be segmented into four categories.

### *Managerial Self-Interest*

Firm managers have limited ability to diversify their own personal wealth position, associated with stock holdings and the capitalization of their career earnings from their own employment position. Therefore, managers prefer stability to volatility because, other things equal, such stability improves their own utility, at little or no expense to other stakeholders.

### *The Nonlinearity of Taxes*

With a nonproportional tax structure, income smoothing reduces the effective tax rate and, therefore, the tax burden shouldered by the firm. By reducing the effective long-term average tax rate, activities that reduce the volatility in reported earnings will enhance shareholder value.[2]

### *The Cost of Financial Distress*

To the extent that the bankruptcy state—or any set of specific states—is associated with a discrete increase in costs, the firm will be forced to recognize this fact in its

[2] Evidence on this point for the insurance industry is provided by Cummins and Grace (1994) and Lamm-Tennant and Rollins (1994).

choice calculus. In such cases, the firm behaves as if it had a concave objective function, because its payoff structure is nonlinear across states.[3]

*The Existence of Capital Market Imperfections*

The existence of the cost imperfections results in underinvestment in some states, where internally-generated funds fall short of the amount of new investment that would be profitable in the absence of these capital market imperfections. The cost of volatility is the forgone investment in each period that the firm is forced to seek external funds. Recognizing this outcome, the firm embarks upon volatility reducing strategies, which have the effect of reducing the variability of earnings.

In each case, the economic decision-maker is shown to face a nonlinear optimization because of the reason offered, and this leads the decision-maker to be concerned with the variability of returns. In the first case, the objective function itself is concave, while in the others the effect of some feature of the economic environment is to lead firm managers to behave in a risk-averse manner.

Together, the stories work fairly well. Firm managers are interested both in expected profitability and the risk, or variability, or reported earnings or market value. The latter can be rationalized by the existence of nonlinear costs across the range of profit states associated with any given expected value. The nonlinearity stems from managerial incentive effects, the tax structure, the costs of crisis, and/or forgone investment opportunities. In any or all of these cases, the firm is led to treat the variability of earnings as a choice variable that it selects, subject to the usual constraints of optimization. How it proceeds to manage the risk position of its activity is the area to which we now turn.

## RISK AS A CENTRAL INGREDIENT IN THE INDUSTRY'S FRANCHISE

*What Risks Are Being Managed?*

Insurers are in the risk business. In the process of providing insurance and other financial services, they assume various kinds of actuarial and financial risks. The risks contained in the insurer's product sales—that is, those embedded in the products offered to customers to protect against actuarial risk—are not all borne directly by the insurer itself. In many instances, the institution will eliminate or mitigate the actuarial and financial risk associated with a transaction by proper business practices; in others, it will shift the risk to other parties through a combination of reinsurance, pricing, and product design. Only those risks that are not eliminated or transferred to others are left to be managed by the firm for its own account. This is the case because the insurance industry recognizes that it should not engage in business in a manner that unnecessarily imposes risk upon it, nor

[3] Staking and Babbel (1995) provide some empirical support for the application of this argument to the insurance industry. The National Association of Insurance Commissioner's risk-based capital requirements have the effect of creating nonlinear costs of financial distress, and insurers are very much aware of this effect (see Cummins, Harrington, and Niehaus, 1994).



should it absorb risks that can be efficiently transferred to other participants. Rather, it should manage risks at the firm level only if they are more efficiently managed there than by the market itself or their owners in their own portfolios. In short, it should accept only those risks that are uniquely a part of the insurer's array of services.

Elsewhere it has been argued that risks facing all financial institutions can be segmented into three separable types from a management perspective (see Oldfield and Santomero, 1997). These are risks that can be eliminated or avoided by standard business practices, risks that can be transferred to other participants, and risks that must be actively managed at the firm level.

In the first of these cases, the practice of risk avoidance involves actions to reduce the chances of idiosyncratic losses from standard insurance activity by eliminating risks that are superfluous to the institution's business purpose. Common risk avoidance practices include at least three types of actions. The standardization of process, insurance policies, contracts, and procedures to prevent inefficient or incorrect financial decisions is the first of these. Another is the construction of portfolios on both sides of the balance sheet that benefit from diversification and the application of the law of large numbers and central limit theorem, which reduce the effects of any one loss experience. Finally, the implementation of incentive-compatible contracts with the institution's management to require that employees be held accountable is the third. In each case, the goal is to rid the firm of risks that are not essential to the financial service provided, or to absorb only an optimal quantity of a particular kind of risk.

There are also some risks that can be eliminated, or at least substantially reduced, through the technique of risk transfer. Markets exist for many of the risks borne by the insurance firm. Actuarial risk can be transferred to reinsurers. Catastrophe risk can be offset somewhat by undertaking a position in catastrophe futures and in catastrophe bonds. Indeed, a number of capital market alternatives for dealing with this kind of risk are currently under consideration (see Jaffee and Russell, 1997). Interest rate risk can be hedged or transferred through interest rate products such as swaps, caps, floors, futures, or other derivative products.[4] Insurance policies and lending documents can be altered to effect a change in their duration and convexity. Equity market risk can be reduced with an appropriate futures position in equities. In addition, an insurer can offer products that absorb some financial risks, while transferring other risks to the purchaser. Defined contribution pension plans and variable universal life policies are clear examples of this approach. Finally, the insurer can buy or sell financial claims and reinsurance to diversify or concentrate the risk that results from servicing its client base. To the extent that the actuarial and financial risks of the insurance policies underwritten by the firm are understood by the market, they can be sold in part or in whole at their fair value. Unless the institution has a comparative advantage in managing these risks, there is no reason for it to absorb such risks rather than transfer them.

[4] For a review of the use of derivatives by insurers, see Cummins, Phillips, and Smith (1997).

However, there are two classes of activities where the risk inherent in the activity must and should be absorbed at the insurance firm level. In these cases, risk management must be aggressive, and good reasons exist for using firm resources to manage insurance-level risk. The first of these includes actuarial exposures where the nature of the embedded risk may be complex and difficult to communicate and transfer to third parties. For example, Progressive Insurance Co. has a definite niche in the high-risk auto insurance business owing to its concentration of underwriting activities, and Lutheran Brotherhood has a natural advantage for writing life insurance to its clientele. A similar situation may arise on the asset side of the business where the insurer holds private placements and other complex, proprietary assets that have thin, or even nonexistent, secondary markets. Communication in such cases may be more difficult or expensive than hedging the underlying risk. Moreover, revealing information about the customer may give competitors an undue advantage. The second case includes risk positions that are central to the insurer's business purpose and are absorbed because they are the *raison d'être* of the firm. Actuarial risk inherent in key insurance lines where the insurer may enjoy a competitive advantage or a market niche is a clear case in point. In all such circumstances, risk is absorbed and must be monitored and managed efficiently by the institution. Only then will the firm systematically achieve its financial performance goal.

### *How Are These Risks Managed?*

What are the necessary procedures that must be in place to carry out adequate risk management for those risks that are essential ingredients to the insurer's franchise? What techniques are employed to both limit and manage the different types of risk, and how are they implemented in each area of risk management?

In general, the management of an insurance firm relies on a variety of techniques in their risk management systems. However, it appears that common practice has evolved such that four elements have become key steps to implementing a broad-based risk management system. These include standards and reports, underwriting authority and limits, investment guidelines or strategies, and incentive contracts and compensation. These tools are established to measure risk exposure, define procedures to manage these exposures, limit exposures to acceptable levels, and encourage decision-makers to manage risk in a manner consistent with the firm's goals and objectives. To see these four parts of basic risk management achieve these ends, we elaborate on each part of the process. Then, we illustrate how these techniques are applied to control each of the specific risks facing the insurance community.

*Standards and reports.* Underwriting standards, risk classification, and standards of review are all traditional tools of risk control. Consistent evaluation and rating of exposures of various types are essential for management to understand the risks on both sides of the balance sheet, and the extent to which these risks must be mitigated or absorbed.

The standardization of financial reporting is the next ingredient. Statutory accounting reports have long been standardized, for better or worse. However, the need here goes beyond public reports and audited statements to the need for management information on actuarial risk, asset quality, and overall risk posture. Such internal reports need similar standardization but much more frequent reporting intervals, with daily, weekly, and monthly reports substituting for the quarterly statutory accounting periodicity.



*Underwriting authority and limits.* A second technique for internal control of active management is the use of position limits and/or minimum underwriting and asset quality standards. In terms of the latter, the domain of risk taking is restricted to only those customers or assets that pass some prespecified quality standard. Then, even for those that are eligible, limits are imposed to cover exposures to counterparties, credits, and overall position concentrations relative to various types of risks. In large organizations, with thousands of positions maintained, accurate and timely reporting is difficult, but even more essential.

*Investment guidelines or strategies.* Investment strategies are outlined in terms of concentration and commitments to particular areas of the market, the extent of desired asset/liability mismatching or exposure to interest rate risk, and the need to hedge against systematic risks of a particular type. These limits lead to passive risk avoidance and/or diversification, because managers generally operate within position limits and prescribed rules. In addition, securitization and even derivative activity are rapidly growing techniques of position management open to participants looking to reduce their exposure to be in line with management's guidelines.

*Incentive schemes.* To the extent that the firm can enter incentive-compatible contracts with senior management, line managers, and sales agents, and make compensation related to the risks borne by these individuals, the need for elaborate and costly controls is lessened. Well designed systems align the goals of managers with other stakeholders in a most desirable way. Most financial debacles can be traced to the absence of incentive compatibility.

## RISKS IN PROVIDING INSURANCE SERVICES

How are these techniques of risk management employed by the insurance sector? To explain this, we must begin by enumerating the risks that the insurance industry has chosen to manage and illustrate how the four-step procedures outlined are applied to risk control in each area.

### *The Actuarial View of Risks*

As a starting point, most of the insurers interviewed classified their risks by adapting a framework that was proposed years ago by the Society of Actuaries' Committee on Valuation and Related Problems. Even though the Society of Actuaries is focused on life insurance and pensions, the property-liability insurers interviewed also had adapted the same risk classification paradigm. The various categories of risks are dubbed C-1, C-2, C-3, and C-4, deriving these names from

the committee assigned to make recommendations on these issues.[5] We begin our review of the perceived risks with an explanation of the industry's own definitions.

*C-1 risks are asset risks,* which arise from the possibility that borrowers of insurer funds may default on their obligations to the company, or that the market value of an insurer's investment assets may decline. They include interest rate risk, credit risk, market risk, and currency risk.

*C-2 risk is pricing risk,* which stems from uncertainty about future operating results relating to items such as investment income, mortality and morbidity, frequency and severity of claims and losses, administrative expenses, sales, and lapses. If an insurer's pricing is based on assumptions that prove inadequate, it may not be able to meet its obligations to policyholders.

*C-3 risk is asset/liability matching risk,* which springs from the impact of fluctuating interest and inflation rates on the values of assets and liabilities. If the impact of fluctuating rates is different on assets than on liabilities, the values of assets and liabilities will change by different amounts, and could expose the insurer to insolvency.

*C-4 risks are miscellaneous risks,* generally thought to be beyond the ability of insurers to predict and manage, but they nevertheless represent real risk to the company. These risks include tax and regulatory changes, product obsolescence, poor training of employees and sales agents, and malfeasance, malversation, or misconduct of managers or other employees. Also included is the risk that laws or legal interpretations will change in a way that will alter the firm's obligations *ex post.* Another manifestation of C-4 risk is that there will be an unforeseen downgrade of acquisitions that could lead to a "run" on the insurer's assets. One firm referred to C-4 risk as "stupidity risk"—failure to employ and retain good people.

Two firms wryly refer to a new category of risk, dubbed C-5 risk, which is the havoc that arises when a person who has strong political ambitions or is running for higher political office is appointed to be state insurance commissioner.

The use of the Society of Actuaries' risk classification taxonomy is viewed merely a useful point of departure by some of the insurance firms we interviewed, while others view it as satisfactory for their purposes. In our view, none of the risk classification schemas we saw was completely satisfactory. However, most of the conceivable risks that would impact insurers are included somewhere on the lists that we saw. In most cases, however, the industry is straining to define the inherent financial risks as part of the C-1 through C-4 paradigm that had been developed years ago. In addition, it appears that most schemas have undue focus on risks in isolation, rather than on their contribution to overall firm risk.

### *The Financial View of Risks*

As an alternative to the actuarial decomposition of risk which is unique to the insurance industry, standard financial risk definitions are increasingly being proposed in the industry. For the sector as a whole, these risks can be broken into six

[5] Our discussion of these risks follows that of Black and Skipper (1994).



generic types: actuarial, systematic, credit, liquidity, operational, and legal risks. Briefly, we discuss each of these risks facing the insurance institution, and we indicate how they are managed. Our focus will be on the financial risks, which include the first four of the risks listed below. Of course, the risks associated with the provision of insurance services differ by the type of service rendered.

*Actuarial risk* is the risk that arises from raising funds via the issuance of insurance policies and other liabilities. It is the risk that the firm is paying too much for the funds it receives or, alternatively, the risk that the firm has received too little for the risks it has agreed to absorb. If an insurer invests these funds in efficiently traded securities, it should expect to have, on average, a zero net economic profit on those securities. Therefore, if the insurer pays too much for its funds, it cannot expect to earn a satisfactory profit in the long run. Another aspect of actuarial risk is that, during any given time period, the underwriting losses will be in excess of those projected. This could happen for two reasons. First, the expectations themselves may be based on an inadequate knowledge of the loss distribution. Second, the losses may exceed their expectations in the normal course of business simply because losses fluctuate around their mean. The degree to which they deviate from the mean will depend, of course, on the characteristics of the loss distribution, which depend on the nature of the risks insured.

*Systematic risk* is the risk of asset and liability value changes associated with systematic factors. It is sometimes referred to as market risk. As such, it can be hedged but cannot be diversified completely away. In fact, systematic risk can be thought of as undiversifiable risk. All investors assume this type of risk whenever assets owned or claims issued can change in value as a result of broad economic factors. Systematic risk comes in many different forms. For the insurance sector, however, three are of greatest concern: variations in the general level of interest rates, basis risk, and (especially for property-liability insurers) inflation.

Because of the insurers' dependence on these systematic factors, most try to estimate the impact of these particular systematic risks on performance, attempt to hedge against them, and thus limit the sensitivity of their financial performance to variation in these undiversifiable factors. To do so, most will both track and manage each of the major systematic risks individually. The first of these is undoubtedly *interest rate risk.* Here, they measure and manage the firm's vulnerability to interest rate variation, even though they cannot do so perfectly. At the same time, insurers with large corporate bond, mortgage, and common stock holdings closely monitor their *basis risk.* Here the concern is that yields on instruments of varying credit quality, liquidity, and maturity do not move together, exposing the insurer to market value variation that is independent of fluctuating liability values. In this case, too, they try to manage, as well as limit, their exposure to it. Finally, to the extent that the frequency and severity of claims are influenced by *inflation risk,* expected losses will also be affected. This is particularly the case where insurance policies are written on a replacement cost basis. The inflation of concern can be general inflation—affecting repair costs, medical costs, and the like—or specific and localized inflation—like the quadrupling of certain building materials costs in

southern Florida shortly after Hurricane Andrew. All three of these systematic risks will be recognized as sources of performance variation.

*Credit risk* is the risk that a borrower will not perform in accordance with its obligations. Credit risk may arise from either an inability or an unwillingness on the part of the borrower to perform in the precommitted contracted manner. This can affect the investor holding the bond or lender of a loan contract, as well as other investors and lenders to the creditor. Therefore, the financial condition of the borrower, as well as the current value of any underlying collateral, are of considerable interest to an insurer who has invested in the bonds or participated in a direct loan.

The real risk from credit is the deviation of portfolio performance from its expected value. Accordingly, credit risk is diversifiable but difficult to eliminate completely, as general default rates themselves exhibit much fluctuation. This is because a portion of the default risk may, in fact, result from the systematic risk outlined above. In addition, the idiosyncratic nature of some portion of these losses remains a problem for creditors in spite of the beneficial effect of diversification on total uncertainty. This is particularly true for insurers who take on highly illiquid assets. In such cases, the credit risk is not easily transferred, and accurate estimates of loss are difficult to estimate.

*Liquidity risk* can best be described as the risk of a funding crisis. While some would include the need to plan for growth, the risk here is more correctly seen as the potential for a funding crisis. Such a situation would inevitably be associated with an unexpected event, such as a large claim or a writedown of assets, a loss of confidence, or a legal crisis. Because insurers operate in markets where they may receive clustered claims due to natural catastrophes, or massive requests for policy withdrawals and surrenders due to changing interest rates, their liabilities can be said to be somewhat liquid. Their assets, however, are sometimes less liquid, particularly where they invest in private placements and real estate. Given this situation, it is important for an insurer to maintain sufficient liquidity to handle easily any demands for cash. Otherwise, an insurer that would be solvent without a sudden demand for cash may have to sell off illiquid assets at concessionary prices, leading to large losses, further demands for cash, and potential insolvency.

*Operational risk* is associated with the problems of accurately processing claims and accurately processing, settling, and taking or making delivery on trades in exchange for cash. It also arises in record keeping, processing system failures, and compliance with various regulations. As such, individual operating problems are small probability events for well-run organizations, but they expose a firm to outcomes that may be quite costly.

*Legal risks* are endemic in financial contracting and are separate from the legal ramifications of credit and operational risks. New statutes, court opinions, and regulations can put formerly well-established transactions into contention even when all parties have previously performed adequately and are fully able to perform in the future. For example, changes in the application of statutes of limitations for filing suits have affected the losses arising from property-liability policies. Similarly, the change to joint and several liability rules has also altered the distribution of risks that may be covered by insurance policies (see Huber, 1988).





Another type of legal risk arises from the activities of an institution's management, employees, and agents. Fraud, violations of regulations or laws, and other actions can lead to catastrophic loss. Even a situation where the insurer legally fulfills all of its contract obligations can result in massive litigation if some policyholders had different expectations or understandings about the performance of their policies than what was specified in the contracts.

Every insurer faces a different exposure to each of these risks, depending on its business mix. In all its activities, an insurer must decide how much business to originate, how much to finance, how much to reinsure, and how much to contract to agents. In so doing, it must weigh both the return and the risk embedded in its asset and liability portfolios. Management must measure the expected profit and evaluate the prudence of the various risks enumerated above to be sure that the result achieves the stated goal of maximizing shareholder value (in the case of a stock insurer) or maximizing ownership interests (in the case of a mutual or reciprocal insurer). If the product's expected profit warrants the risk, then the activity is added to the insurer's balance sheet, and the risk must be managed. This risk management is achieved through the four-step process outlined above. How this is implemented for each of the key financial risks enumerated above is the focus of the next section.

In concluding this section, we observe that the actuarial versus the financial views of risk are worlds apart. This difference is a critical barrier impeding the progress toward effective financial risk management. Although both views encompass most of the same risks, the contexts in which they are measured and weighed give rise to different treatments of risk. The actuarial view, with its focus on risks in isolation and their impact on the statutory accounting statements, tends to foster risk measures that do not aggregate well at the firm level and leads to a piecemeal approach toward risk management. In contrast, the financial view looks at risk only insofar as it impacts firm economic value. Risk aggregation and covariance are the focus. Until the actuarial and financial views of risk converge and a unified view of risk is accepted, there will continue to be some disarray in how financial risk is managed by insurers.

## Insurance Risk Management Systems

### *Actuarial Risk*

The risk of paying too high a price to raise funds is an important risk, particularly in light of the fact that insurers raise few funds in the competitive capital market. Most of their debt is raised in the form of issuing insurance policies. Policies are written today in exchange for lump sum or periodic premiums, but the amounts and timing of the repayment of these funds are often unknown and may occur within a month or more than 80 years later. Because the pricing of the policies reflects not only expected losses but also the yields an insurer can earn on the funds between the inception of a policy and its termination or the payment of benefits,

the interest assumption used in developing insurance prices is of critical importance. Two things complicate this process. Forward interest rates cannot be synthesized to lock in a spread, for the insurer has no way of knowing if future periodic premium payments will be forthcoming. Also, the loss distributions can undergo substantial evolution over time, as more information is revealed and as the economic environment changes.

Insurers typically are quite skilled in managing actuarial risk. The manner in which this is done is described in insurance and actuarial textbooks (see Black and Skipper, 1994; Beard, Pentikäinen, and Pesonen, 1984; Gerber, 1979; and Cummins and Derrig, 1989). Therefore, here we will focus on what developments have occurred during the past decade that improve an insurer's ability to price and manage this risk.

*Life insurers.* Until recently, life insurance prices were developed using conservative static assumptions regarding loss distributions and interest rates. Although this approach was satisfactory for much of the past century, it was ill-equipped to accommodate the interest rate volatility that began during the late 1970s. Life insurance policies are replete with options—settlement options, policy loan options, over-depositing privileges, and surrender or renewal privileges on the part of the insured, and discretionary dividend and crediting rate options on the part of the insurer. Indeed, some have even viewed a life insurance policy as little more than a package of options (see Smith, 1982). In stable interest rate environments, policyholder utilization of these options is often predicated on individual or family circumstances. Hence, in the aggregate, utilization rates are fairly steady and amenable to forecasting.

However, when interest rates are volatile, the options gain in value and their utilization rates can fluctuate wildly. Traditional actuarial methods, which depended upon stability, were incapable of correctly valuing these options; hence, many policies were woefully underpriced.[6] Today, the standard valuation methods that have been adopted by most of the sophisticated life insurers explicitly value these embedded options. Thus, insurers now can estimate the cost of the various option-like provisions of all kinds of life insurance policies. Most life insurers we interviewed were using the PTS software of Chalke, Inc. (now SS&C, Inc.). This software, and competing software offered by Tillinghast and others, uses modern stochastic valuation techniques, familiar in the pricing of fixed income and mortgage-backed securities, to estimate the values of insurance policies in a manner consistent with that used to value the assets. Needless to say, this represents a big advance in the tools with which insurers can practice risk management.

Lest our enthusiasm for this advance be misconstrued as euphoria, we hasten to add that all is not well here. First, the stochastic valuation methodology most

[6] The value of the policy loan option by itself could account for 20 to 45 percent of the present value of all future insurance premiums, if the option were used optimally. When factoring in the suboptimal utilization of this option, the estimated cost to an insurer of providing this option was in the 8 to 12 percent range. Yet insurers had historically charged *nothing* for this option. Indeed, it was simply mandated that insurers begin to offer this option.



commonly used relies on a single stochastic factor. Most fixed income and mortgage-backed security valuation models are based on at least two stochastic factors. Without two factors, one tends to produce model values that are too highly correlated and whose movements in value are perfectly correlated. Also, the speed of the software is sufficiently slow that it is difficult to implement more than a handful of path simulations in arriving at "option-adjusted" values. Moreover, it is unrealistic to attempt to model a prepayment feature or a call feature, which may be triggered by changes in long-term yields, while using the short-term rate paths to value the instrument. The second drawback is perhaps even more serious. Most insurers have inadequate data collected and assembled with which to reliably model the interest sensitivity of policy option utilization. Accordingly, the valuation models really allow an insurer only to quantify better the impact of its guesses about what those utilization functions might look like. We encountered much frustration among life insurers that, even though the valuation software had taken a long time to develop, the data requirements of the valuation software have still not been met. The third drawback is in how insurers interpret the data analyses provided by "black boxes." We found that, in some companies, there is neither an understanding of nor an appreciation for the risk measures produced. There is often such shallow understanding of the underpinnings of the methodologies employed to measure risk that the computer output is either disregarded or uncritically accepted. Over time, as insurer personnel receive more training in stochastic methods and the meaning of the risk measures, they will be able to use the software to greater advantage in measuring and managing their risks.

Nonetheless, the availability of valuation software that is consistent with modern valuation principles is an important step forward, and software that is currently under development will remedy the shortcomings of being based on a single stochastic factor and producing value estimates, dare we say, at a relaxed pace. With this software, actuaries currently produce pricing estimates based on a dozen or so scenarios. However, they typically also test their prices using hundreds and thousands of additional scenarios, albeit not in an option-adjusted framework.

*Property-liability insurers.* Financial risk management is closely related to the issue of financial pricing, especially for property-liability insurers.[7] In the property-liability insurance sector, there is no counterpart to the modern valuation software for pricing liabilities. However, option-adjusted arbitrage-free valuation tools may be overly powerful given the imprecision associated with many of the risks that are insured. There are few options that compare with those available in life policies, and nominal values are often not guaranteed; guarantees are sometimes in terms of covering repair costs, replacement costs, medical costs, and so forth. Even when there are nominal maximum amounts of coverage, the losses below the maximum are subject to additional uncertainty because of inflation.

The use of reports and standards for underwriting life/health and property-liability risks is routine. It is common to have dozens, and sometimes over a hun-

[7] See Cummins and Harrington (1987), Taylor (1994), and Derrig (1994) for a discussion of this important related issue.

dred, "cells" in which to classify the risks. Base rates can be related to a number of factors, such as age, gender, occupation, schooling, health status and history, property characteristics, nature of business, and so forth. These base rates are then adjusted to reflect experience factors (e.g., past claims, driving behavior). While the fair premiums will be a function of interest rates, in practice the premiums charged will not adjust to reflect current interest rates very often. This is because it is administratively cumbersome to alter insurance premium schedules every time the interest rates change.

Underwriting limits are commonly established. Authority is limited to a certain amount. While insurance agents may have temporary binding authority, it is a common practice to have a party who is not involved in the policy sale to review the underwriting and make a determination whether the risk will be ultimately accepted and insured. Insurers are typically better at keeping track of sales commissions than at tracking losses to a particular sales agent or underwriter. However, many of the leading life/health and property-liability insurers are carefully tracking the experience of their sales and underwriting personnel. If the experience falls outside the norm, it is common to place restrictions on further sales or more severe limitations on underwriting; alternatively, the activities of these sales agents and underwriters could be subject to greater oversight.

Perhaps the area of greatest concern in the area of actuarial risk is the misalignment of incentives between owners of the insurance firm and its sales and marketing staff. Much can be done to improve it. The typical arrangement is to pay commissions for sales of new policies, with the commissions on a multiperiod contract heavily front-loaded, particularly for life/health products. This creates a tremendous incentive for agents to sell as much business as possible, whether it is profitable for the company or not. It also creates strong incentives to replace existing policies, whose commission rates have dwindled to the low single digit percentage range, with new policies that pay commissions ranging from 20 to 100 percent of first-year premiums. Sales managers and marketing personnel are also often rewarded based on volume of sales. Even senior management may sometimes have their compensation tied to sales growth.

Experience has shown that rapid growth is one of the factors most commonly associated with insolvency. It is useful to remember here that what is growing most rapidly is the accumulation of liabilities, not assets. One way to foster rapid growth is to underprice liabilities. Employees and agents whose compensation is tied to sales growth are therefore strong proponents of more "competitively-priced" insurance policies. Senior management often comes from a sales background and is sympathetic with the notion that what is good for the insurance agents is good for the company. Pricing actuaries, who are supposed to be the watchdogs and gatekeepers in this area, are often placed under tremendous pressure to alter their assumptions so that the company's products can be priced more competitively. Of course, over time it will become apparent if the insurance policies are mispriced, but that is weighed against the immediate benefits of higher commission earnings and growth.



The sales side has one powerful club in this battle for determining policy prices. Sales agents often work for a number of insurers and can shift new business toward them. Worse, they can take existing business away from the firm, before it breaks even from heavy initial policy costs, and direct it elsewhere if they can demonstrate satisfactorily that policy illustrations or prices appear to be more favorable elsewhere. Many firms in the insurance industry are well aware of this misalignment of interests, yet feel thwarted by regulations about commission schedules.

In the long run, of course, insurers offering noneconomic policies will go bankrupt. But the long run can take a long time to arrive; hence, the insurer who is trying to rationally price its policies faces a quandary. Does it succumb to the uneconomic pricing temporarily and hope to survive beyond the irrational players, and then restore sensible pricing, or does it choose to write very little current business and lose its distribution force? Neither choice is an attractive alternative.

*Systematic Risk*

*Systematic risk of liabilities.* No area in financial risk management of insurance has evolved as much as the analysis of systematic risk of liabilities during the past decade. This is, in large measure, due to the fact that insurers feel an increased sense of urgency in applying the tools of asset/liability management to measure and manage interest rate risk. We note that the two most recent Goldman Sachs surveys of life insurance chief investment officers ranked asset/liability management at the top of the list of their concerns, whereas the topic did not surface in the top four rankings in their earlier surveys. Property-liability insurers are also giving greater attention to the area.

The increased importance given to asset/liability management is echoed in the 1995 and 1996 surveys of Lamm-Tennant and Lamm-Tennant and Gattis, who found it near the top of the factors that influence investment policy. These findings are notable because they combine the results of both life/health and property-liability insurers, and cover companies that are much smaller than those in the Goldman Sachs surveys. When contrasted with the earlier surveys of Babbel and Lamm-Tennant (1987), Babbel and Klock (1988), Lamm-Tennant (1989), and Bouyoucos and Siegel (1992), the increased importance of interest rate risk and asset/liability management during the past few years is remarkable. All of the life/health and property-liability insurers we interviewed perceive this source of risk to be crucial to understand, measure, and manage. However, the insurers we interviewed run the gamut from naive to very sophisticated when it comes to measuring interest rate risk.

On the liability side of the balance sheet, most of the life insurers were using PTS software developed by Chalke, Inc. to measure the effective duration and convexity of their liabilities.[8] The others were using TAS from Tillinghast or some

[8] Measures of interest rate sensitivity that take into account the interest-sensitive cash flows of an asset or liability stream are referred to as "effective duration and convexity" or, alternatively, "option-adjusted duration and convexity." Measures of interest rate sensitivity which [illegible]me all cash

internally-developed software. Most of the life insurers who are using the commercially-available software packages have implemented some of their own customized enhancements to meet better their needs, capabilities, focus, and concerns.

The use of effective duration and convexity measures represents a quantum leap from what the practice was only a few years ago. Prior to 1992, virtually none of the insurers had access to a commercially-available software package that could compute measures of effective duration and convexity for their liabilities. Even the PTS, TAS (formerly CALMS), and Milliman and Robertson software packages available at that time would not produce measures of effective duration and convexity. Rather, the duration numbers, in those cases where they were produced, were simple modified or Macaulay measures, which assume that all cash flows are fixed. Yet, liabilities are virtually all interest sensitive to some degree. The traditional duration measures produced errors so large as to lead to reckless investment decisions, while imbuing such decisions with a veneer of analytical and quantitative credibility.[9] Back then, insurers who were concerned about interest rate risk relied heavily upon simulations. Indeed, duration estimations were considered so primitive that they were generally eschewed in favor of simulations, and rightfully so in our opinion. This is because many of the duration estimates that we saw then did not fully incorporate the interest rate sensitivity of cash flows for either assets or liabilities.

Today, convexity measures are also produced by the PTS software that is most commonly used. We found that insurers place less confidence in the convexity numbers produced than in the duration numbers. This is because convexity numbers are much more sensitive to lapse assumptions than are duration numbers; while a misspecification of the interest rate sensitivity of lapses and other options can cause a large error in effective duration estimates, it will cause an even greater error in the estimates of convexity. Most insurers feel that they do not have enough reliable data on which to specify the relation of lapses and policy surrenders to interest rate movements. The lack of confidence they have in this crucial input to convexity estimates translates into a lack of confidence in the convexity estimates themselves. However, most companies do pay some attention to convexity estimates but place wide ranges around those estimates. The most common way to grasp the impact of convexity is in toggling the lapse/surrender sensitivity parameters in numerous simulations. The standard among the companies we interviewed is to perform simulations of between 500 and 10,000 paths to capture the impact of changing interest rate levels on policy lapses/surrenders.

While life insurers have more interest than confidence in the convexity estimates, they have progressed a long way over the past few years. Prior to 1992, the

---

flows are fixed, or at least insensitive to movements in interest rates, include "modified duration and convexity" and "Macaulay duration."

[9] For example, we estimated the duration on a block of participating whole life policies for one mutual company. Its Macaulay duration was around 22 years, while its effective duration was approximately 5.6. See Lamm-Tennant (1989) for a revealing survey of the level of sophistication in understanding and applying duration and Babbel (1994b) for a discussion of the pitfalls in using the older duration measures.



commercially-available software did not even produce convexity estimates for life insurance liabilities. Instead, firms relied almost entirely on simulations. Many firms used only the seven highly artificial scenarios required of New York's Regulation 126. Prior to its passage, some insurers did not use the simulation method at all. Rather, they relied simply on their "best point estimates" and static lapse assumptions. Even today, there are insurers who use nothing more than the seven scenarios required under Regulation 126 to assess their exposure to interest rate risk.

The property-liability insurers with whom we spoke have less concern about interest rate risk than their life/health counterparts. Nonetheless, they manifest greater understanding of the problem than they had a few years earlier. All of them are well aware of the importance of measuring duration of assets, producing in-house estimates of duration, or acquiring them from outside vendors, for most of their fixed income assets. More problematic is the estimation of duration of their real estate and equity portfolios. However, an analysis of the duration of their liabilities is generally missing. They have a notion that the duration is relatively short—perhaps a couple of years or so—but no more specific information (see Babbel and Klock, 1993). Convexity is even less of a concern for these property-liability insurers. Nonetheless, there appears to be at least some interest rate sensitivity in the payments made to satisfy property-liability claims. For example, it is well known that workers' compensation claims tend to increase during periods of unemployment, as fraudulent claims are filed with greater frequency. Similarly, fires and arson tend to occur with greater frequency when insured values exceed market values. To the extent that these and other situations are linked to interest rate levels, it can be supposed that some property-liability liabilities are interest sensitive. Available evidence on this front is scant, however (see Choi, 1992). Where the sensitivity is measurable, it tends to be more closely linked to inflation than to nominal interest rates. Therefore, the influence of inflation on their liabilities is deemed more important.

It is fair to say that most property-liability insurers pay little attention to the duration of their liabilities. It is generally thought that interest rate risk accounts for only a small portion of the change in the value of liabilities over time and that other risks, such as actuarial risk, price regulation, legal risk, underwriting risk, inflation risk, and event or catastrophe risk, swamp the influence of interest rate movements on the pricing and valuation of property-liability insurance liabilities.

*Systematic risk of assets.* Insurers are concerned with interest rate risk more than other systematic risk factors, and rightly so. Over the past two decades, it has been the source of much of the fluctuation in the value of fixed income assets, which constitute the majority of their assets. However, while it is the crucial systematic risk on the life insurance liability side, and of some importance for property-liability liabilities as well, it is prominent but less dominating on the asset side of the balance sheet. This is because asset values are perceived to be affected not only by general interest rate levels, but also by basis risk, default risk, liquidity risk, call risk, prepayment risk, extension risk, sinking fund options, convertibility, real estate, and equity risk. Yet, several of these risks are simply different manifes-

tations of interest rate risk, making accurate measurement of paramount importance.

The measurement of interest rate risk on the asset side of the balance sheet is generally well done, although some insurers have a long way to go. Many insurers use the actuarial software mentioned above to estimate the durations and convexities of their investments. Some use software and pricing services like GAT and Bloomberg that are oriented strictly toward the asset side of the balance sheet. Several have developed their own, more sophisticated in-house programs for estimating values of both sides of the balance sheet. It is common to use more than a single source to assess the duration and convexity of assets. One stated reason for this is the divergence of opinion between the various programs and pricing services.

We did not encounter any property-liability insurers who carefully measure the interest rate risk of both sides of their balance sheet. It is more common for the focus to be on the interest rate risk of only the assets. Here the tools of duration and convexity measurement are applied, and insurers take steps to manage the overall exposure of their assets to interest rate risk and keep it within some targeted range. It is common for property-liability insurers to use interest rate futures, swaps, and options to moderate this risk to acceptable levels. Options and futures are also used to hedge equity market risks, where the insurer maintains a large position in common stocks. The hedges are put in place, and then removed, as market conditions change and the insurers' appetites for equity risk wax and wane.

*Asset/liability management.* Asset/liability management typically does not go far beyond an assessment of the impact of interest rate movements on the value of the firm. Other systematic risks are usually dealt with in a more piecemeal fashion. The standard practice is to produce estimates of liability durations and convexities for each line of business, as well as for each asset class. These estimates are then weighted by the fair value of liabilities, or market value of assets, to arrive at overall asset and liability duration and convexity estimates. After factoring in leverage, the insurers are able to obtain measures of surplus duration and convexity. Examples of product level and firm level analyses are given in Tables 1 and 2.

The frequency for providing analysis of interest rate risk varies widely. Some firms provide weekly summaries of their asset durations and convexities, and monthly or quarterly summaries of their liabilities. In the case of interest rate futures and options, reports are more frequent, owing to their tremendous impact on overall interest rate risk. Some firms assess their liability interest rate risks on only an annual basis, and among the property-liability companies, liability durations are often not measured at all, nor is the interest rate sensitivity taken into account in liability simulations.



**Table 1**
**Option-Adjusted Duration Convexity for Life Insurance Product XX**

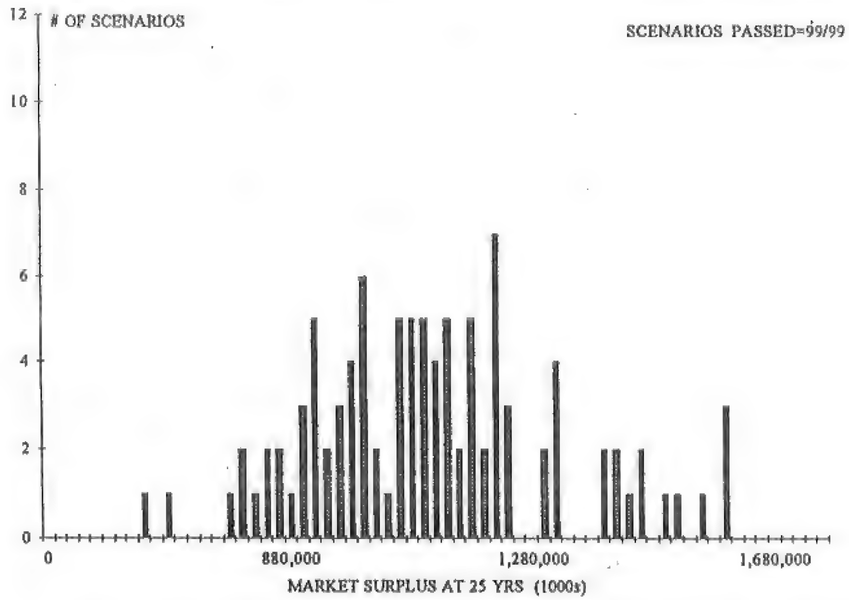
| | Assets | | | Liabilities | | | Surplus | | |
|---|---|---|---|---|---|---|---|---|---|
| Interest Rate Shift | Option-Adjusted Value | Effective Duration | Convexity | Option-Adjusted Value | Effective Duration | Convexity | Option-Adjusted Value | Effective Duration | Convexity |
| −200 | 1600 | 2.9 | N.A. | 1545 | 4.1 | N.A. | 55 | −30.8 | N.A. |
| −150 | 1500 | 3.0 | −66 | 1405 | 2.9 | 60 | 95 | 4.5 | −3000 |
| −100 | 1400 | 3.1 | −85 | 1330 | 2.1 | 450 | 70 | 22.1 | −25 |
| −50 | 1300 | 3.3 | −176 | 1243 | 1.8 | 700 | 57 | 36.0 | −125 |
| 0 | 1200 | 3.5 | −190 | 1151 | 1.5 | 680 | 49 | 50.5 | −10000 |
| 50 | 1100 | 3.8 | −121 | 1061 | 1.4 | 390 | 39 | 69.1 | −8,555 |
| 100 | 1000 | 4.2 | −50 | 975 | 1.2 | 85 | 25 | 121.2 | −6000 |
| 150 | 900 | 5.0 | 0 | 901 | 1.1 | 20 | −1 | −3508.9 | 100 |
| 200 | 800 | 5.9 | N.A. | 845 | 1.0 | N.A. | −45 | −86.1 | N.A. |

**Table 2**
**XYZ Insurance Company Surplus Duration Analysis**

| Asset Class | Market Value | Effective Duration | Liabilities | Fair Value | Effective Duration | Economic Surplus | Effective Duration |
|---|---|---|---|---|---|---|---|
| Bonds | 10,000,000 | 6.5 | SPDA 1 | 8,500,000 | 3.2 | | |
| Mortgage-Backed | 5,000,000 | 4.2 | SPDA 2 | 4,000,000 | 2.0 | | |
| Preferred Stock | 200,000 | 8.1 | Universal Life | 400,000 | 7.9 | | |
| Common Stock | 3,000,000 | 2.1 | Term Life | 300,000 | 1.3 | | |
| Mortgage Loans | 6,000,000 | 5.7 | Whole Life | 12,000,000 | 6.8 | | |
| Equity Real Estate | 2,000,000 | 1.0 | Endowment | 500,000 | 8.5 | | |
| Short-Term and Cash | 1,500,000 | 0.8 | | | | | |
| Total | 27,700,000 | 4.74 | | 25,700,000 | 4.85 | 2,000,000 | 3.36 |

Many companies couple this kind of analysis with one that shows the distribution of the future market value or, more typically, book value of surplus, based on hundreds of scenarios. An example is provided in Figure 1. This approach is conceptually fine, although we caution that, when looking at distant future values of surplus, the values produced are extremely sensitive to slight variations in assumed yield spreads, which can get compounded for 30 years and often are overly optimistic. Rarely are these approaches implemented with sufficient skill to account for the various correlations and patterns that can be observed in practice.

**Figure 1**
**Distribution of Surplus Accumulation**
**Results of 99 Random Scenarios**




Although many firms use the same general frameworks for analysis, when it comes to implementation we begin to see divergence in the quality of inputs and practice. By relying on a number of outside sources to provide the estimates of interest rate sensitivity for assets and liabilities, a number of insurers have injected another risk into the mix: divergent technologies and assumptions. We believe that, for the purposes of asset/liability management, it is a misdirected effort to obtain the most credible measures of interest sensitivity of certain assets or liabilities. It is far more important to get measures of interest rate sensitivity that are calibrated similarly. After all, the precise values are of less importance here than their relative values and the implication these have for the volatility of equity.

We saw a number of practices that invite problems. One prominent insurer uses effective duration and convexity estimates for the liability side of the balance sheet and modified Macaulay duration measures for the asset side. Another insurer does exactly the opposite. Some insurers base their aggregate duration and convexity numbers on book value weights, rather than market value weights. More than one insurer is frustrated that its actuarial department relies entirely on simulations and provides no duration or convexity measures whatsoever. The actuarial scenarios are based on completely random interest rate paths, inconsistent with any financial theory or history of interest rates. Several insurers rely on liability duration estimates based on only one interest rate factor, but on asset duration estimates based on two factors. Some insurers use duration measures for corporate bonds



and mortgage-backed securities supplied by Wall Street that are based on different volatility parameters and processes than those used for other asset and liability categories. Some estimates are based on lattice models, while others are based on simulation models, or simulating through lattices. Only one insurer we know of is attempting to correct the duration measures on corporate bonds for the basis risk between corporate and government bonds.[10] Some insurers take into account the basis risk between movements in long-term vs. short-term interest rates, but many do not. Some take into account all kinds of potential twists in the yield curve, while others allow only for parallel shifts.

In setting limits on the amount of systematic risk the company desires to retain, a common approach is the one used by a leading multiline insurer. The company places limits on its desired portfolio structure to reflect the variety of risks to which it is exposed. Limits are set on individual asset holdings, on industry concentration, and on asset type—including mortgage-backed securities and collateralized mortgage obligations—all in a risk-based capital context. However, nowhere did we observe a methodology to derive such limits, or even a standardized approach across business lines.

For the balance sheet as a whole, limits are employed in two different ways. One approach is to impose a limit on the amount of duration mismatch allowed, either for particular product lines or for aggregating across all assets and liabilities. For instance, one company applies these restrictions on a product segmentation basis, allowing up to a year duration mismatch on participating whole life products, but only one-tenth of a year on GICs. Another company does not place restrictions on duration mismatches on a product-by-product basis, but on an aggregate portfolio basis. In our view, although most companies we interviewed use some sort of product segmentation approach, it is not necessary to do so. The advantage of a segmentation approach is the discipline it imposes on the pricing process, so that long-term yields are not used for pricing short-term liabilities, and so forth. However, if this same discipline can be achieved in the pricing of insurance policies without a segmentation of assets into various product groupings, it seems that advantage would disappear. On the other hand, valuable resources would not be consumed in notionally dividing up the general account into the various segments, and it could be managed on an aggregate basis. This would avoid the costly duplication of efforts, where one product manager is selling an asset and another is buying the same or a similar asset, incurring transaction costs. Some firms simply transfer the assets between portfolio segments and use some sort of internal transfer pricing mechanism. Assets are acquired by the firm and then allocated to each product group according to perceived needs. This is done to foster a better sense of accountability and used in performance evaluation. But to reward a product group for producing net profits between liability costs and rates of return on assets which they had no responsibility in acquiring or divesting seems to be rewarding them for risks over which they had no control. While we appreciate the

[10] See Babbel, Merrill, and Panning (1987) for an explanation of the correction procedure.

need for pricing discipline and control, we feel this could be achieved more simply and that the asset portfolios can be managed better on an aggregate basis.

The other limit is a restriction on the amount of scenarios that are allowed to reveal losses due to asset/liability mismatches. These limits typically are placed not only on the distribution of final simulated results, but also on the evolution of solvency over time associated with the simulations. One leading firm has almost no tolerance for scenarios showing negative profitability due to interest rate risk exposure. Because it is persuaded that interest rates are virtually impossible to forecast, and over which it has no control, it has decided to avoid interest rate risk of any kind, to the extent possible.

At the firm that decided to avoid as completely as possible interest rate risk, portfolio managers are not rewarded in any way for taking interest rate risk and trying to "time" the market. Indeed, their job could be lost if they stray outside narrow boundaries. Some firms purport to eschew interest rate risk, yet reward their investment department personnel if they achieve investment income or total rates of return above some benchmark level. By measuring only periodically the interest rate risk of assets, this invites the portfolio managers to "game" the system and attempt to improve their returns by incurring interest rate risk for brief periods of time. Some firms have duration targets but ignore convexity, leading portfolio managers to try barbell, ladder, or bullet maturity approaches to achieve higher investment income, depending on the shape of the term structure.

But by in large, the major difference between investment practices that we saw during this study, compared to what was occurring less than a decade ago, is that there is far less emphasis on yield and more on total rate of return. As recently as five years ago a survey of the American Council of Life Insurance revealed that two-thirds of chief investment officers did not even consider total rate of return as an investment objective. Yield was the primary focus. This was an impediment to effective asset/liability management but is beginning to dwindle. Nonetheless, we observed more concern with book yield than we feel is appropriate, given its lack of importance to the true economic performance of the firm.

*Credit Risk*

In addition to the credit risk that reveals itself as basis risk in the systematic risk factors listed above, there is also the risk of default on significant firm investments. Although it may be idiosyncratic risk to the market as a whole, it is not idiosyncratic risk to the insurer maintaining a significant position in an asset that goes into default.

Insurance firms are generally very focused on credit risk, as are rating agencies and regulatory authorities. They produce weekly and monthly reports that monitor the credit risk of their assets. They rely on outside rating agencies, such as Moody's, Standard and Poor's, Duff and Phelps, and Dunn and Bradstreet. In addition, virtually all of their investments are assigned credit ratings by the Securities Valuation Office of the National Association of Insurance Commissioners, which are used for statutory reporting purposes. These ratings are not always viewed as sufficient measures of credit risk for those insurers who feel that absorbing credit risk is an important part of their franchise. Many insurers have their own due diligence requirements to meet before they will take on an investment that has credit risk. They undertake internal credit risk ratings, in some ways quite similar to those of Moody's or Standard and Poor's, although with different weightings on the risk factors. Moreover, they are prone to update their internal credit risk ratings promptly as important information bearing on the creditworthiness of a major investment position is revealed.



Insurers produce "watch lists" of firms they feel are in financial jeopardy, likely to be downgraded or become insolvent. They often have a dual track credit risk assessment, one for the asset itself, and another for the underlying collateral. They place limits on the portfolio exposure by industry, geographic region, business (e.g., real estate prohibited), and company. They also have lists of approved counterparties for brokerage, settlement, and swaps.

Tables 3 and 4 provide an example of one set of investment guidelines, with general and specific authorizations and limits that we feel are representative of the industry. However, we have seen substantial variation in the practices. Perhaps the poorest approach we saw was a firm that uses Moody's ratings and assigns numerical values to each rating class. For instance, 1 is assigned to a rating of Aaa, 2 is assigned to Aa, 3 to A, 4 to Baa, 5 to Ba, 6 to B, 7 to Caa, 8 to Ca, 9 to C, and 10 to D. Adjustments are made to accommodate the modifiers of 1, 2, and 3 that Moody's often uses to designate relative quality within a ranking class. The company then has a target number of 3 to achieve in its overall credit risk plan. One problem with this approach, which they recognize, is that default rates and volatility of default rates do not grow linearly as rating is decreased step by step. Coupled with an incentive structure that rewards portfolio managers for the investment yields they book, this system leads to a credit barbell approach, as shown in Figure 2, because the portfolio manager can achieve superior yields by doing so.

The best approach we saw included a more refined ranking of credit risk, not by letter but by default probability coupled with standard deviation of defaults for each ranking. Covariance of asset returns is also taken into account, and the entire credit risk problem is cast in a surplus-oriented mean-variance model. Diversification guidelines are incorporated through constraints on the optimization. Liquidity risk is reflected by reducing the expected returns by a number of basis points deemed appropriate from historical experience.

**Table 3**
**Investment Guidelines**
**General Authorizations**

These general authorizations are to remain in effect until January 1997 unless modified or canceled.

The total investment in any one credit (total of bonds, preferred stock, convertible securities and common stock) is limited to 1 percent of net admitted assets, with the exception of direct U.S. Treasury and full faith and credit obligations and U.S. government-sponsored enterprise obligations, as further specified below.

The following authorizations specify which transactions the investment officers of the company are authorized to conduct at their discretion.

*Fixed Income Securities—Nonconvertible*

A. Purchase U.S. Treasury and full faith and credit obligations in unlimited amounts.
B. Purchase U.S. government-sponsored enterprise obligations. Such purchases to be limited to 3 percent of net admitted assets per enterprise and an aggregate limitation of 10 percent.
C. Purchase corporate, municipal, or foreign bonds denominated in U.S. dollars which are rated in the Baa category or better; purchase private placement issues rated in the Baa category or better, or its equivalent.
D. Purchase mortgage-backed securities and collateralized mortgage obligations consistent with investment committee limitations and the specific authorizations.
E. Purchase asset-backed securities consistent with investment committee limitations and specific authorizations.
F. Purchase preferred stocks of companies whose bonds or preferred stocks are rated in the Baa category or better, or its equivalent.

*Short-Term Fixed Income*

H. Purchase commercial paper (with maturities not exceeding 270 days), certificates of deposit, or bankers acceptances which are rated either A-1 or better by Standard & Poor's or P-1 by Moody's. Such purchases to be limited to 1 percent of net admitted assets per credit.

*Repurchase Agreements*

I. Repurchase agreements may be purchased with banks or security dealers as designated in the specific authorizations with the following characteristics: at least 102 percent collateralized by U.S. Treasury or agency obligations, for periods not exceeding 60 days. Total amount outstanding limited overall to 5 percent of net admitted assets and with any one counterparty to 2 percent.

**(Continued)**



**(Table 3, Continued)**

*Reverse Repurchase Agreements*

J. Reverse repurchase agreements may be transacted with banks or security dealers by specific authorization. Total amount outstanding to be limited to 5 percent of net admitted assets and per counterparty to 2 percent.

*Dollar Rolls*

K. Enter into dollar roll transactions with banks or security dealers designated in the specific authorizations. Total amount outstanding to be limited to 5 percent of net admitted assets and per counterparty to 2 percent.

*Other*

L. Enter into CMO residual commitments as specified in the specific authorizations.
M. Use financial futures contracts and interest rate options (exchange traded and over-the-counter) to reduce interest rate risk exposure.
N. Sell securities held in the portfolio. Any purchaser or transfer agent of such a security need not inquire into the authority for such sale upon the secretary's certification that it is made under this subdivision "N."

*Convertible Securities*

O. Under the limits specified for nonconvertible bonds, purchase convertible bonds of companies with issues rated in the Baa category or better or of companies with lower or nonrated issues according to the following schedule:

| Moody's (or S&P) Rating | Authorized Limit per Issuer as Percentage of the Convertible Portfolio (%) |
|---|---|
| Ba1 (BB+) to Ba3 (BB-) | 2.0 |
| B1 (B+) to B3 (B-) | 1.0 |
| No Rating | 1.0 |
| Caa (CCC+) to C (D) | 0.5 |

Purchases must not cause the aggregate statement value of convertible debt rated less than 2 by the NAIC to exceed the following percentages of the convertible debt portfolio:

| NAIC Rating | Restrictions (%) |
|---|---|
| 3 3Z 6 6Z | 60.0 |
| 4 4Z 6 6Z | 40.0 |
| 5 5Z 6 6Z | 7.5 |
| 6 6Z | 2.5 |

**(Continued)**

**(Table 3, Continued)**

P. Purchase of medium or lower grade bonds must not cause the aggregate statement value of bonds, including convertible debt, rated less than 2 by the NAIC to exceed the following percentages of net admitted assets:

| NAIC Rating | Restrictions (%) |
|---|---|
| 3 3Z 6 6Z | 10.0 |
| 4 4Z 6 6Z | 5.0 |
| 5 5Z 6 6Z | 3.0 |
| 6 6Z | 1.0 |

Q. No more than 25 percent of the portfolio's bonds rated less than 2 by the NAIC can be companies within a single industry, nor should these issues in aggregate conduct business within one narrow geographic region. The duration of these issues as a group should be viewed in the context of the total bond portfolio.

R. Purchase convertible preferred stocks under the limits specified for nonconvertible with issues rated Baa or better or of companies with lower or nonrated issues according to the following schedule:

| Moody's (or S&P) Rating | Authorized Limit per Issuer as Percentage of the Convertible Portfolio (%) |
|---|---|
| Ba1 (BB+) to Ba3 (BB-) | 2.0 |
| B1 (B+) to B3 (B-) | 1.0 |
| No Rating | 1.0 |
| Caa (CCC+) to C (D) | 0.5 |

*Common Stock*

S. The investment officers are authorized to purchase common stocks of any U.S. or Canadian corporation or any foreign corporation whose shares are included in the S&P Common Stock Index in amounts consistent with the investment policy statements as approved by the committee subject to these additional restrictions. Specifically, the investment officers may not:

1. Invest in the equity securities of closed-end funds, investment companies, limited partnerships, or real estate investment trusts without specific authorizations.
2. Make any purchase of common stock which would result in more than 5 percent of the value of the common stock portfolio being invested in the securities of one issuer.
3. Purchase a common stock if, as a result, more than 25 percent of the assets of the common stock portfolio will be invested in a particular industry.

**(Continued)**



**(Table 3, Continued)**

4. Purchase a common stock if, as a result, more than the authorized limit of the assets of the common stock portfolio will be invested in common stock of that particular issuer. The authorized limit per issuer will be a function of the issuer's common stock market capitalization in accordance with the following schedule:

| Market Capitalization of Issuer of Common Stock | Authorized Limit per Issuer as a Percentage of the Overall Common Stock Portfolio (%) |
|---|---|
| $0 - $25 million | 0 |
| $26 - $50 million | 0.25 |
| $51 - $100 million | 0.50 |
| $101 - $200 million | 1.00 |
| $201 - $500 million | 2.00 |
| $501 - $1,000 million | 3.00 |
| $1,001 - $2,000 million | 4.00 |
| More than $2,000 million | 5.00 |

Furthermore, the investment officers will make a quarterly presentation to the investment committee on the performance of the common stock portfolio and its risk characteristics in relation to appropriate benchmarks.

T. The investment officers are authorized to use stock index futures and options contracts (exchange traded and over-the-counter) to reduce stock market risk exposure.

*Real Estate/Commercial Mortgages*

U. The Investment Officers are authorized to:

1. Commercial Mortgages and Real Estate Equity
Consummate transactions for the purchase, sale, exchange, and disposition of real estate loans or equities, up to $2.5 million with the approval of any three of the following officers of the company: chairman of the board, president, chief investment officer and senior real estate investment officer. All transactions completed under this authority will be reported to the real estate subcommittee of the investment committee of the board of directors

**Table 4**
**Investment Guidelines**
**Specific Authorizations**

The following specific authorizations for the purchase of securities to be executed at the discretion of the officers of the company were renewed by the committee:

1. *Bonds, Short-Term Investments, Convertibles Subordinated Debentures*
   None
2. *Common Stocks*
   A. *Domestic*
      None

| | *Authorization Percent* |
|---|---|
| B. *Foreign Companies* | |
| DeBeers Consolidated Mines (ADR) | 1.0% |
| Sea Containers Ltd. | 1.0% |

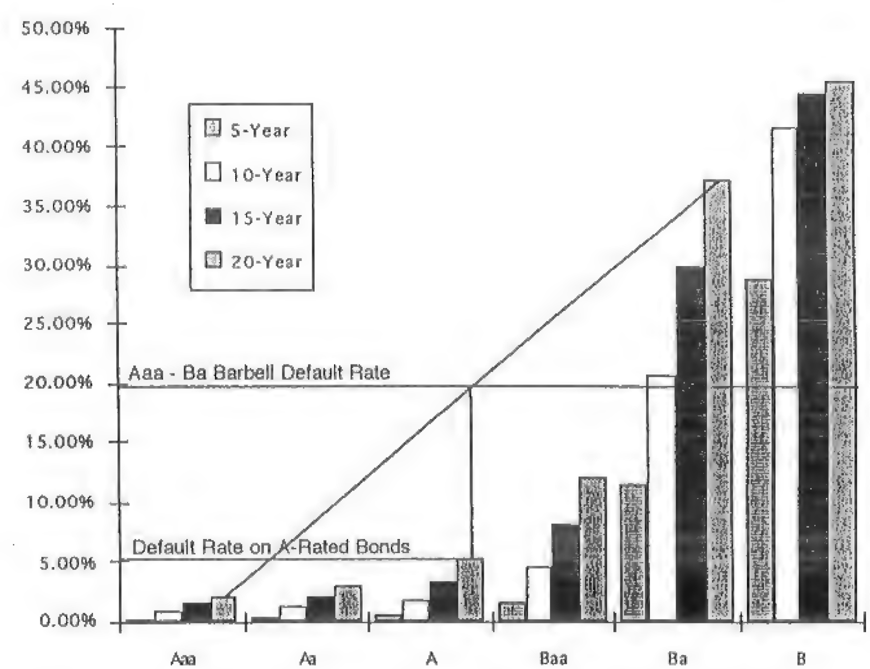
3. *Convertible Preferred Stock*

| | |
|---|---|
| Sea Containers Ltd. | 1.0% |

4. *Repurchase Agreements, Reverse Repurchase Agreements and Dollar Roll Transactions*
   There is a list of banks and security dealers authorized for repurchase agreements and reverse repurchase agreements and dollar roll transactions.
5. *Hedging Transactions*
   A. The following officers of the investment department are hereby authorized to execute hedging transactions in accordance with the guidelines:
      Executive Vice President & Chief Investment Officer
      Senior Vice President, Vice President, and Second Vice President, Equity Securities
      Senior Vice President, Vice President, and Second Vice President, Fixed Income
      Assistant Vice President, Equity Trader
   B. There is a list of banks and brokerage firms approved for the establishment of futures trading accounts.
   C. Maximum amount which may be hedged is 5 percent of admitted assets.
6. *Residual Commitments*
   The maximum amount of CMO Residual commitments is limited to $100 million.
7. *Mortgage-Backed Securities Commitments*
   In the aggregate, the maximum amount of FHLMC, FNMA, and whole loan commitments is limited to $750 million at book value.
8. *Asset-Backed Securities*
   In the aggregate, the maximum amount of ABS is limited to $250 million at book value.
9. *Short-Term Borrowing*
   Up to a maximum amount of $100 million, such maximum amount not to exceed $50 million with any one bank or financial institution



**Figure 2**
**5-, 10-, 15-, and 20-Year Cumulative Default Rates, 1970-1995**




*Liquidity Risk*

Although insurance companies face liquidity risk, most of the insurers we interviewed have little concern for it. Only one is concerned about having too little liquidity, and one is concerned about having too much liquidity. The others do not seem to be concerned, believing their situation to be well managed.

*Life insurers.* Liquidity is not as big a concern with many insurance firms as it is in other financial institutions for one good reason: most of their policies are less liquid than their assets. Life insurers issue policies that commonly feature high surrender charges. These charges are either explicitly stated or implicit in the schedule of cash buildup. For example, a single premium deferred annuity, with annual crediting rate reset and a seven-year maturity, may feature surrender charges beginning at 7 to 10 percent during the first year, and declining in steps toward zero at maturity. Similarly, universal life and whole life products often have very low surrender values during the first year or two of a policy, and begin building up rapidly after that point. Other policies, such as variable universal life, have surrender provisions that act much like a mutual fund, where the amount received depends on the value of the underlying fund.

A decade or so ago, the problem of illiquidity was more prononounced. Policy surrenders for some companies approached a 60 percent rate, and massive amounts of policy loans were withdrawn. The industry has managed this problem

in two ways. First, much of the new business written is sensitive to market rates of interest, so that there is not as wide a divergence between crediting rates on life policies and annuities vs. market rates of interest. Thus, the incentive to surrender a policy is lessened. Second, policy loans are now mostly offered at variable rates that track market rates of interest, rather than the fixed policy loan rates of yore. An alternative to the variable rates charged on loans is a process known as "direct recognition," whereby the schedule of cash value buildup is altered if policy loans are incurred. The tax environment has also changed, and interest paid on policy loans is now no longer a deductible expense for tax purposes. Thus, the interest rate arbitrage incentive has virtually disappeared.

Not long ago, however, several well known life insurance companies, such as First Capital Life, Fidelity Bankers Life, Executive Life of California, Executive Life of New York, and Mutual Benefit, experienced severe liquidity problems. But in each of these cases, there were other factors that precipitated a "run on the bank" phenomenon. The run in each case was caused, in part, by well publicized investment performance problems. In the case of the first four of the above mentioned companies, there were vast sums of policies in force with minimal, and even zero, surrender charges. Some of the policies had been marketed through Wall Street brokerages and were therefore of the "hot money" variety. Although the first four of these companies had large amounts of liquid assets, the withdrawal rates were so high that even these liquid resources were strained. Generally, however, life insurers are managed in such a way as to avoid these runs on the bank, either through policy design, sales channel, investment policy, or level of surplus.

Another problem that was more prevalent a decade ago was that all bonds were essentially reported at amortized book values. If a bond was sold prior to maturity, any capital gain or loss would need to be recognized. Many insurers, particularly during the early 1980s, could not afford to have the large capital losses appear on their books and impact their surplus. Today, however, a large portion of fixed income holdings are placed in accounts available for sale or trading, and are therefore already marked to market. Therefore, whether they are sold does not impact the reported values or the surplus as much.

*Property-liability insurers.* On the property-liability side of the business, the liquidity risk comes mostly from event risk or catastrophe risk. Most policies do not feature cash values that are easily accessible through surrender, although some policies will allow insureds to cancel prior to maturity and receive a portion of their premium. Moreover, the policies are typically very short term, renewable annually. With the relatively large surplus positions of most property-liability insurers, most policyholders are willing to ride out a storm, knowing that renewal time will approach in just a few months.

In terms of the accounting policies of property-liability companies, they tend to mark-to-market a larger portion of their fixed income securities than life insurers and also tend to hold more liquid, shorter-term securities on average. Thus, their liquidity concerns almost always stem from event risk or catastrophe risk. They can avoid some of these risks by reinsuring portions of their books of business and by broadly diversifying their risk portfolios geographically, by industry, and by



type of risk. Alternatively, they can hold large amounts of liquid surplus assets. One leading property-liability insurer attempts to keep enough surplus on hand to accommodate a "once in 500 years" event.

We provide some tables that are representative of the practice of liquidity risk measurement and management. The company defines liquidity risk as the risk of having inadequate net cash flows to meet expenses, benefits, withdrawals, and loan payments. It views product liquidity risk as the fluctuations in cash flows outside of the ranges that are expected. It first ranks assets by relative liquidity. It then projects its cash flows over a multiple year horizon on both sides of the balance sheet using a form like Table 5. For those firms that maintain segmented accounts by product line, these liquidity reports are generated by product, as in the case of Table 6.

A company's asset/liability committee is typically responsible for measuring and managing liquidity risk. Despite the lack of concern regarding this risk, there is still a large amount of analysis that is done to guard against illiquidity. In this regard, liquidity risk decisions are part of the analytics and scenario testing used by the company. In its investment plan, management of liquidity risk is two-fold. First, the company uses corporate and Regulation 126 modeling to measure net cash flow under various interest rate scenarios. Second, control is achieved by imposing constraints on investment. One such constraint includes ensuring that over 50 percent of the assets are held in "marketable securities."

Many companies use PTS for most of their analysis and stress testing. The scenario testing includes about 50–100 scenarios which shift the yield curve, both parallel and slope changes. For each path created by the scenario, net cash flows must be positive. Solvency of the company is also determined along each path. Reports are used for each product, and cash flows are projected out for about 30 years under each of the scenarios. By aggregating across products for each scenario, the company has an idea of the distribution of liquidity at the firm level.

After all the scenario tests, the results have little impact on immediate decisions. For instance, if net cash flows were negative for a large portion of the stress tests, this would not imply that asset composition would change immediately. Suggestions would of course be made, but there is no guarantee that the portfolio would change immediately. Similarly, concentrations are not altered unless there are modifications to limits which would only occur quarterly or annually.

**Table 5**
**Liquidity Report**
**Life Insurance Product XXX Projected Cash Flows**

| Period | Asset Cash Flow | Liability Cash Flow | Net Cash Flow |
|---|---|---|---|
| 1996 | | | |
| 1997 | | | |
| 1998 | | | |
| . | | | |
| . | | | |
| . | | | |
| 2015 | | | |

In addition to running these scenario tests, there is also a "worst case scenario." This includes cash outflows of over 300 percent over a 2 to 3 year period and a lapse rate of 45 percent. For comparison, the highest lapse rate the company has ever experienced is 18 percent, when it decided to decrease significantly dividends after a long history of either increasing or maintaining dividend payments.

There are no managers at the companies studied whose performance is based on the management of liquidity risk. Although it would be difficult to base everyday compensation on unlikely events, the company may benefit in the future from having a clearer line of responsibility with regard to liquidity management. The general procedure is that any ongoing problems with liquidity would be brought to the attention of the chief financial officer. Any changes to the company's credit ratings which could potentially affect liquidity demands are very much a concern of the auditing group.

### *Other Risks Considered But Not Modeled*

Beyond the basic four financial risks—actuarial, systematic, credit, and liquidity—insurers have a host of other concerns, as indicated above. Some of these, like operating risk, are a natural outgrowth of their business, and insurers employ standard risk avoidance techniques to mitigate them. Standard business judgment is used in this area to measure the costs and benefits of both risk reduction expenditures and system designs, as well as operational redundancy. While generally referred to as risk management, this activity is substantially different than the management of financial risk addressed here.

Yet there are still other risks, somewhat more amorphous, but no less important. In this latter category are legal, regulatory, reputational, and environmental risk. In each of these risk areas substantial time and resources are devoted to protecting the firm's franchise value from erosion (see Babbel, 1994a). As these risks are less amenable to *a priori* financial measurement, they are generally not addressed in any formal, structured way. However, they are not ignored at the senior management level of the insurance firm.



**Table 6**
**Relative Liquidity of Assets and Liabilities for Life Product XXX**

| | Current Quarter | Prior Quarter −1 | Prior Quarter −2 | Prior Quarter −3 |
|---|---|---|---|---|
| *Liabilities*[a] | | | | |
| Level 1 Liquidity | | | | |
| Level 2 Liquidity | | | | |
| Level 3 Liquidity | | | | |
| Level 4 Liquidity | | | | |
| (Levels 1 + 2) ÷ Total | | | | |
| *Assets* | | | | |
| Level 1 Liquidity | | | | |
| Level 2 Liquidity | | | | |
| Level 3 Liquidity | | | | |
| Level 4 Liquidity | | | | |
| (Levels 1 + 2) ÷ Total | | | | |
| *Summary* | | | | |
| Level 1 (Assets) ÷ Level 1 (Liabilities) | | | | |
| Level 1–2 (Assets) ÷ Level 1–2 (Liabilities) | | | | |
| Guideline | | | | |
| Level 1–3 (Assets) ÷ Level 1–3 (Liabilities) | | | | |

| Note: | Level 1 | Level 2 | Level 3 | Level 4 |
|---|---|---|---|---|
| Liabilities | Book | Int Grnty Expires | Market; | Now Allowed |
| Surrender is at: | Surcharge < 2 Percent | + Scheduled | Market and Surcharge; | |
| | | Ann Paymts, | Book, | |
| | | Next 12 months | Surcharge > 2 Percent | |
| Assets | Cash, STI Inv Grade Bonds with aggregate market loss < IMR; Other | Mtg amort NII over Nest 12 months | Inv Grade Bonds with aggregate market loss in excess of IMR | Mortgage; Below Investment Grade Bonds; Affil; VC: Other |

[a] Liabilities are net of policy loans.

In passing from this topic it is worthwhile and timely to pause to consider one of the legal risks now encroaching upon the life insurance industry. During the course of our interviews, a number of firms had been sued in the area of misrepresentation of insurance products by insurance agents. At the time of this writing, 64 class action lawsuits had been filed against firms for their so-called "vanishing premium" policies whose premiums did not vanish as illustrated, owing to a prolonged decline in market interest rate levels. New class action lawsuits were being filed at the rate of one every three days. The damages claimed are staggering for some of the companies.

The manner in which insurers are responding to these lawsuits ranges from attempts to gain a quick and comprehensive settlement to attempts to have the arguments heard in court. Some insurers are merely biding their time to see how other firms fare in the struggle. But more interesting is how insurers are acting to avoid future problems stemming from alleged agent misrepresentation. One of the leading firms has established a large department of compliance to train and monitor the behavior of its sales agents. All of its sales and promotional literature is undergoing careful scrutiny by the legal department.

Another firm has created an auditing division to oversee compliance from a central location, computerizing each transaction. Management is concerned with five components of compliance: customer satisfaction, new products, stable earnings, expansion capabilities, and corporate miscellaneous. They are building controls into the centralized computer system. Should an agent exceed the allotted number of address changes, disbursements, lapses, or sales, the computer will not process the policy until the auditing department has had a chance to investigate further. These stop measures are not announced to either the customer or the sales agent. The director of auditing indicated that the system is intended to prevent problems rather than react to them.

A compliance division has been introduced to complement the role of the internal audit group. This division is responsible for insuring the field force and also providing training to sales agents so they will better be able to represent the company's products. This division is currently sending out surveys to its customers to find out if they really understand the products they now hold. It is hoped that these measures will mitigate any class action suits in the future.

## AREAS WHERE FURTHER WORK WILL IMPROVE METHODOLOGY

Thus far, the techniques used to measure, report, limit, and manage individual risks have been presented. In each of these cases, a process has been developed, or at least has evolved, to measure the risk considered, and techniques have been deployed to control each of them.

The insurance industry is clearly evolving to a higher level of risk management techniques and approaches than had been in place in the past. Yet, as this review indicates, there is significant room for improvement. Before the areas of potential value added are reviewed, however, it is worthwhile to reiterate an earlier point. The risk management techniques reviewed here are not the average, but the techniques used by firms at the higher end of the market. The risk management approaches at smaller institutions, as well as larger but relatively less sophisticated ones, are less precise and significantly less analytical. In some cases, they would need substantial upgrading to reach the level reported here. Accordingly, our review should be viewed as a glimpse at best practice, not average practices. Nonetheless, the techniques employed by those that define the industry standard could use some improvement. By category, recommended areas where additional analytic work would be desirable are listed below.





### *Actuarial Risk*

There remains too much disagreement in the most fundamental area of actuarial science—namely, what discount rate or rates does one use to value insurance liabilities (see Altman and Vanderhoof, 1996). With such broad disagreement about what insurance is worth, or what it will cost the insurer, it is little wonder that we encounter difficulties when it comes to managing risk. On the bright side, it must be acknowledged that there is a flurry of activity taking place directed toward solving this conundrum. In our opinion, as the scope of what exactly is being valued is more carefully defined, there is a convergence in the estimates produced by the alternative valuation methods (see Babbel, 1997).

During the past decade, tools have been developed that can take into account the interest rate sensitivity of policy cash flows. However, many insurers have not employed these tools. Among those who have, there is a severe problem with the data inputs that are necessary to produce useful output. Insurers have not tracked or organized their lapse, surrender, and claims data in a manner that allows them to accurately model their interest rate sensitivity. Although the models are capable of accommodating virtually any functional form of this behavior—including the effects of policy seasoning, channel of distribution, and so forth—little data exist to estimate the functional form. Of course, this was also the case with regard to the modeling of mortgage-backed securities prepayment a decade ago, but since that time data were collected and analyzed, allowing for the enlightened application of the valuation tools that we see today in that sector. We expect that the same will be true for insurers in a few more years.

Another area where we expect to see rapid improvement is in the versatility of the actuarial software and its speed. Currently, only single-factor stochastic models are being widely used. This results in output that is inconsistent with the other side of the balance sheet, where two or more factors are typically used. Moreover, the speed of processing is so slow that insurers make undesirable compromises when it comes to modeling their products fully. Emerging computer technology undoubtedly will remove this impediment to better policy pricing and risk analysis.

### *Systematic Risk*

Tremendous progress has already been made over the past five years in managing systematic risk. Most of it has been directed toward interest rate risk management, which is appropriate given its importance to insurers. An important area for fur-

ther development is the incorporation of basis risk and equity risk. Another important advance will be a consistent valuation methodology for both sides of the balance sheet.

While simulation studies have substantially improved over the past few years, the use of book value accounting measures and cash flow losses continues to be problematic. Movements to improve this methodology will require increased emphasis on market-based accounting. Such a reporting mechanism must be employed on both sides of the balance sheet, however, not just the asset portfolio.

The simulations also need to incorporate the advances in dynamic hedging that are used in complex fixed income pricing models. As it stands, these simulations tend to be rather simplistic, and scenario testing is rather limited.

### *Credit Risk*

The evaluation of credit rating continues to be an imprecise process. We note divergence between the NAIC ratings assigned to particular public and private placements vs. the ratings assigned by the Wall Street ratings agencies. We should never expect to see a complete convergence here, as there is no single set of weights to apply to the risk factors across all industries and firms. However, we do expect to see less divergence over time, as more becomes known about the factors that lead to default.

We also expect to see more enlightened practices when it comes to aggregating credit risks. A sensible aggregation scheme would take into account default rates, default losses, and the shape of the distribution of losses across all ratings categories. In time, we may even see a move toward market-based default measures, at least on publicly-traded debt instruments. We anticipate that credit risks will soon be evaluated in a framework consistent with other financial risks. Some insurers are already moving in this direction.

### *Liquidity Risk*

For the life insurers, liquidity risk seems to be the least of their major financial risks. Most companies are doing a satisfactory job of managing this risk. With the advent of mark-to-market accounting, the problems for liquidity caused by the fiction of book accounting will gradually subside. Most life insurers model this risk well. In the property-liability sector, it remains a large risk. Crisis models need to be linked better to operational details. In addition, the usefulness of such exercises is limited by the realism of the environment considered.

If liquidity is to be managed, the price of illiquidity must be defined and built into illiquid positions. While this logic has been adopted by some institutions, this pricing of liquidity is not commonplace.

### *Risk Aggregation and Knowledge of Total Firm Exposure*

The quest for an estimate of aggregate firm risk has been a stumbling block for the insurance industry. The extent of the differences across risks of different types is quite striking. Actuarial risk is carefully modeled but reported at infrequent intervals. There is often a lack of follow-up to see whether, based on the insurer's experience, the actuarial assumptions have been appropriate. Systematic risk, particularly interest rate risk, is typically measured by life insurers on both sides of the balance sheet and by property-liability insurers at least on the asset side. Interest rate risk exposure is discerned using measures of effective duration and convexity, scenario simulations, or a combination of the two. For assets it may be reported as often as weekly or monthly, but for liabilities it is generally reported only quarterly or annually. The credit risk process is a qualitative review of the performance potential of different bonds and borrowers. It results in a rating, periodic reevaluation at reasonable intervals through time, and ongoing monitoring of various types or measures of exposure. Liquidity risk, on the other hand, more often than not, is dealt with as a planning exercise, although some reasonable work is done to analyze the effect of adverse events that affect the firm.



## CONCLUSION

The analytical approaches that are subsumed in each of these analyses are complex, difficult, and not easily communicated to nonspecialists in the risk considered. The insurer, however, must select appropriate levels for each risk and select, or at least articulate, an appropriate level of risk for the organization as a whole. How can and is this achieved?

The simple answer is "not very well." Senior management often is presented with myriad reports on individual exposures, such as specific credits, and complex summaries of the individual risks, outlined above. The risks are not dimensioned in similar ways, and management's technical expertise to appreciate the true nature of both the risks themselves and the analyses conducted to illustrate the insurer's exposure to them is limited. Accordingly, over time, the managers of specific risks have gained increased authority and autonomy. In light of recent losses, however, things are beginning to change.

At the organizational level, overall risk management is being centralized into a risk management committee, headed by someone designated as the senior risk manager. The purpose of this institutional response is to empower one individual, or group, with the responsibility to evaluate overall firm-level risk, and determine the best interest of the company as a whole. At the same time, this group is holding line officers more accountable for the risks under their control and the performance of the institution in that risk area. Activity and sales incentives are being replaced by performance compensation which is based not on business volume, but on overall profitability.

At the analytical level, aggregate risk exposure is receiving increased scrutiny. To do so, however, requires the summation of the different types of risks outlined above. This is accomplished in two distinct but related ways. In the first approach, risk is measured in terms of variability of outcome. Where possible, a frequency distribution of net returns is estimated from historical data, and the standard deviation of this distribution is estimated. Capital is allocated to activities as a function of this risk or volatility measure. Then, the risky position is required to carry an expected rate of return on allocated capital, which compensates the firm

for the associated incremental risk. By dimensioning all risk in terms of loss distributions, and allocating capital by the volatility of the proposed activity, risk is aggregated and priced in one and the same exercise.

The second approach is similar to the first, but depends less on a capital allocation scheme and more on cash flow or earnings effects of the implied risky position. This approach can be used to analyze total firm-level risk in a similar manner to the first approach. Again, a frequency distribution of net returns from any one type of risk can be estimated from historical data. Extreme outcomes can then be estimated from the tail of the distribution. Either a worst-case historical example is used for this purpose, or a three or four standard deviation outcome is considered. Given the downside outcome associated with any risk position, the firm restricts its exposure so that, in the worst-case scenario, the insurer does not lose more than a certain percentage of its surplus or current income. Therefore, rather than moving from volatility of equity value through capital, this approach goes directly to the current earnings implications from a risky position. The approach, however, has two very obvious shortcomings. It is cash flow based, rather than market value driven; and it does not necessarily directly measure the total variability of potential outcomes through *a priori* distribution specification. Rather, it depends upon a subjectively prespecified range of the risky environments to drive the worst-case scenario.

Both measures attempt to treat the issue of tradeoffs among risks using a common methodology to transform the specific risks to firm-level exposure. In addition, both can examine the correlation of different risks and the extent to which they can or should be viewed as offsetting. As a practical matter, however, only two of the insurers interviewed that were using these approaches viewed the array of risks as a standard portfolio problem. Rather, they separately evaluate each risk and aggregate total exposure by simple addition. As a result, much is lost in the aggregation. Perhaps over time this crucial issue will be addressed more widely.

The ability of insurance companies to estimate and manage firm-level risk is a long way off. To reach this goal requires much more precision in the estimation and management of the individual risks within the firm. Aggregation has meaning only to the extent that the individual elements can be aggregated. This presumes that they are measured correctly, dimensioned in a similar manner, and incorporated in a unified framework of risk. When this is accomplished, risks of different types will be contrasted and compared, and tradeoffs will become possible. However, to achieve this requires a significant amount of work on the individual risks within the industry before any reasonable aggregation can transpire.



## References


Alexander, J. N., P. J. Bouyoucos, M. H. Siegel, and A. Gola, 1994, The Goldman Sachs Insurer CIO Survey, *Industry Resource Group* (New York: Goldman Sachs).

Altman, E. I. and I. T. Vanderhoof, eds., 1996, *Fair Value of Insurance Liabilities* (Homewood, Ill.: Irwin).

Babbel, D. F., 1994a, A Perspective on Model Investment Laws for Insurers, *C.L.U. Journal*, 48: 72–77.

Babbel, D. F., 1994b, Asset-Liability Matching in the Life Insurance Industry, in: E. Altman and I. T. Vanderhoof, eds., *The Financial Dynamics of the Insurance Industry* (Homewood, Ill.: Irwin).

Babbel, D. F., 1997, Financial Valuation of Insurance Liabilities, in: E. Altman and I. T. Vanderhoof, eds., *Fair Value of Insurance Liabilities* (Homewood, Ill.: Irwin).

Babbel, D. F. and D. R. Klock, 1988, Insurance Pedagogy: Executive Opinions and Priorities, *Journal of Risk and Insurance*, 55: 701–712.

Babbel, D. F. and D. R. Klock, 1993, Measuring the Interest Rate Sensitivity of Property/Casualty Insurer Liabilities, in: S. G. Gustavson and S. E. Harrington, eds., *Insurance, Risk Management, and Public Policy* (Boston: Kluwer Academic Publishers).

Babbel, D. F. and J. Lamm-Tennant, 1987, Trends in Asset/Liability Management for Life Insurers, *Insurance Perspectives* (New York: Goldman Sachs).

Babbel, D. F., C. B. Merrill, and W. Panning, 1997, Default Risk and the Effective Duration of Bonds *Financial Analysts Journal*, 53: 35–44.

Barton, E., G. Bodnar, and A. Kaul, 1994, Exchange Rate Variability and the Riskiness of U.S. Multinational Firms: Evidence from the Breakdown of the Bretton Woods System, Working Paper 94-6, Wharton Weiss Center, Philadelphia.

Beard, R. E., T. Pentikäinen, and E. Pesonen, 1984, *Risk Theory*, Third Edition (New York: Chapman and Hall).

Black, K., Jr. and H. D. Skipper, Jr., 1994, *Life Insurance*, Twelfth Edition (Englewood Cliffs, N.J.: Prentice-Hall).

Bouyoucos, P. J. and M. H. Siegel, 1992, The Goldman Sachs Insurer CIO Survey, *Industry Resource Group* (New York: Goldman Sachs).

Choi, J. H., 1992, Estimating Market Value of Loss Payments of Property/Liability Insurance Companies Using Modern Valuation Technology, Working Paper, Wharton School, Philadelphia.

Cummins, J. D. and R. A. Derrig, 1989, *Financial Models of Insurance Solvency* (Boston: Kluwer Academic Publishers).

Cummins, J. D. and G. Grace, 1994, Tax Management and Investment Strategies of Property-Liability Insurers, *Journal of Banking and Finance*, 18: 12–28.

Cummins, J. D. and Scott Harrington, 1987, Fair Rate of Return in Property-Liability Insurance (Norwell, Mass.: Kluwer Academic Publishers).

Cummins, J. D., Scott Harrington, and Greg Niehaus, 1994, Risk-Based Capital Requirements for Property-Liability Insurers: A Financial Analysis, in: Edward Altman and Irwin Vanderhoof, eds., *The Financial Dynamics of the Insurance Industry* (Homewood, Ill.: Irwin).

Cummins, J. David, Richard D. Phillips, and Stephen D. Smith, 1997, Corporate Hedging in the Insurance Industry: The Use of Financial Derivatives by U.S. Insurers, *North American Actuarial Journal*, 1: 13–49.

Derrig, Richard A., 1994, Theoretical Considerations of the Effect of Federal Income Taxes on Investment Income in Property-Liability Ratemaking, *Journal of Risk and Insurance*, 61: 691–709.

# Know Your Derivatives Before You

by Roger A. Haynes

WHARTON REPROGRAPHICS

*To prevent catastrophic losses from derivatives trading, it is important for risk managers to make sure their organizations understand the risks and establish the proper controls.*

Lurking in the treasury department of a corporation somewhere sits a 28-year-old MBA, eyes glued to a computer screen, risking the future of the company as he tweaks a financial model and selects an interest rate or currency swap. He may have been manipulated by an investment banker or broker, too self-assured to contradict, but he just may have bet the company's pension assets on a miscalculated 500-basis-point rate swing.

The huge derivatives trading losses that brought down the British merchant bank Barings PLC is just one example of this risk management nightmare. Consider the West Coast insurance company that fell into liquidation because it bet the wrong way on an interest rate trend. Or the municipal finance manager, under pressure to keep taxes low with revved-up investment returns, who cost taxpayers millions by taking risks he either was misled into or didn't understand. Or the team of commodity traders that developed a strategy to hedge its core energy product price. The traders got into a disagreement with their bank over the strategy's risk level and funding requirements. The dispute caused huge losses and a revocation of the company's credit lines.

These are all examples of traders, treasury staff and operational people taking what they believe to be prudent business risks. But lately it seems that more and more derivatives transactions are making headlines as they unravel and rebound to sting their companies. In the end, shareholders may take aim at management and unleash breach of fiduciary duty or negligence actions that allege mismanagement on a grand scale. Directors' and officers' liability suits could rain on management as the highly automated plaintiffs bar swings into action following a 10 percent drop in the company's share price.

In this environment, underwriters, management and the press have forced members of the risk management community to ask about the risk profile of their company's derivatives and hedging activity. Sometimes risk managers are fearful of suits alleging inappropriate speculation with shareholders' money. Other times they may test internal controls designed to protect a company's viability in the face of management error or fraud. Prudent risk managers, together with brokers and underwriters, are making proactive efforts to identify the scope of this risk and are establishing control measures and policy provisions to treat it.

The notional amount of all derivatives—financial instruments deriving their value from some underlying asset, such as currency, interest rates or commodities—is estimated to exceed $15 trillion and, in fact, may be as high as $35 trillion. Corporate managers are using these arrangements at a staggering pace and, they and their advisors are quick to assure us, are doing so for good, sound and carefully considered reasons. Depending on the point of view, derivatives are either conservative protec-

*Roger A. Haynes is managing director at Sedgwick Financial Risk Specialists in Boston, MA.*

# Risks Leap

"derivatives" was virtually unheard of outside treasury departments. Hedging activities within operating units had not been seen as an appropriate subject to discuss with the risk manager, even though interest rate hedges and insurance rates are nearly indistinguishable in their key components. Insurance is nothing other than a hedge—some might say a particularly inefficient hedge, but a hedge nonetheless.

Part of the problem rests with traditional organizational boundaries. Most treasury departments have dealt in a different realm from the risk managers, brokers and underwriters who design and implement insurance-based hedge techniques for a company. Risk managers' awareness of the issues surrounding trading in the treasury department and how

tion devices or speculative gambles. One thing seems clear—they are here to stay. A survey conducted by *Institutional Investor* in late 1994 estimated that as few as 3 percent of corporate treasurers expected to decrease their derivatives activity in 1995, even in the wake of several well-publicized debacles.

#### Growing Awareness

As recently as a year ago, there were still companies, municipalities and other entities with their collective heads in the sand. The term

losses could affect the company overall has generally been equally low. In fact, as the term "risk management" is used in treasury operations, it often has little or nothing to do with traditional risk and insurance management. Instead, the term generally refers to getting advice and designing models to control and monitor a strategy designed by investment bankers and dealers.

There are myriad risks faced by companies using derivatives instruments:

***Market and price-setting risk***—interest rate

PAUL COLON

volatility, currency fluctuations and other risks. The variability within these risks is further compounded by the skill level of the people that negotiate the trade.

***Counterparty risk***—the credit risk that promises made under these agreements may not be fulfilled.

***Control risk***—the adequacy of internal measures to establish trading limits and monitor the enthusiasm of managers caught in the heat of a deal.

***Communication risk***—the effectiveness of dealers and advisors in telling management the scope of the risk in the hedge effort. This also applies to management's ability to convey those issues to shareholders.

***Legal risk***—particularly in the cross-border trading arena, where derivatives may be considered gambling or where the enforceability of international settlement agreements is in question.

Each of these risks can be treated by traditional risk control methods, but often the responsibility for the construction, implementation and monitoring of the appropriate control technique rests in different hands. This can make coordination across business units and among levels of management difficult.

Traditionally, the role of risk managers in derivatives trading comes after the fact—usually with a phone call from management asking about insurance protection to recover from a loss. Typical questions are asked about the trading coverage in a fidelity bond, wrongful act definitions in a D&O policy or ERISA provisions covering imprudent investments in a pension portfolio. Depending upon the specific circumstances, coverage and indemnity depends on the answers to these important questions.

### COMMUNICATE AND CONTROL

Effective risk management begins with clear communications about the exposures a company faces and the controls it has established. In the derivatives arena, this has taken on several colors. Bankers and traders have designed proprietary financial models that assist in meeting current requirements calculating the market value of a company's positions on a frequent, even daily, basis. These models are becoming essential management tools now that the Financial Accounting Standards Board (FASB) has required disclosure of hedging positions in the financial statements of companies whose fiscal years ended after December 15, 1994. Seeing net positions marked to market in the financial statement notes has given shareholders, risk managers, underwriters and brokers a clear view of management's aggressiveness and skill in this arena—and enabled them to monitor a company's enthusiasm for speculation. A core issue for indemnification concerns often rests with determining whether the actions taken with shareholder, pensioner or policyholder funds represent prudent hedging or speculation. Some recommended risk management practices for derivatives are described in Table 1.

But even more fundamental issues call for attention in this realm. Someone in authority has to understand the risks and potential losses. If you are unsure about the scope of the risks, bring in an investment banker or a consultant to review your exposures and hedging techniques. Because some people might say there could be a potential conflict of interest if someone offering strategic advice also sells derivatives products, it is important to ensure that the disclosure is complete and that the integrity of the advisor is unquestionable.

The Banker's Trust agreement with the New York Federal Reserve Bank, reached to settle allegations stemming from derivatives losses at Gibson Greetings, Inc., gives guidelines on marketing and communications that all but demand the client understand what he or she is buying. Many dealers are providing daily

**Table 1**

In 1993, the Group of Thirty, an international study group of economists and bankers, released recommended risk management practices for derivatives monitoring and use that fall into several main areas.

**The board of directors should review and approve all derivatives activity, either directly or through designated committees or specific senior managers. The board should also review and approve all changes to trading and hedge strategies.**

**Written policies and procedures must be in place to control derivatives use. These policies should specifically delineate the practices allowed and determine who is responsible for controlling the risks.**

**Monitoring functions need to be formally established to test and report the status of the hedge and its various potential outcomes and to provide timely reports of current positions.**

**Formal audits of a company's derivatives procedures and positions must be performed by outside professionals.**

**Immediate reporting of changes in value (basically calling for daily mark-to-market capability) is required in a company's financial reports.**

valuations of their clients' derivatives positions. In this environment, "transparency" has become the new byword. As communications improve between derivatives dealers and users, and companies better understand what they are getting into, knowledgeable risk taking is more likely than in the past.

As in any other area, it should be fundamental that risk managers understand the corporation's tolerance for financial loss. Logically, the same guidelines should apply in the conduct of a hedge strategy as in establishing insurance retentions, deductibles and limits. Unfortunately, most companies make these decisions independently of one another, without establishing consistent guidelines or calculating aggregate loss exposures.

Knowing the cost of a hedge in relation to the assumed risk is also fundamental. Risk charges on interest rate hedges are often one or two basis points. Compared to rates for insurance products, this may seem reasonable—until the potential for loss is factored into the equation. Unlike insurance transactions, where potential losses are usually limited to the premium cost, large gains as well as large losses are possible in derivatives trades.

It is also important to understand any deal and to avoid getting carried away with the upside potential. Speculation for profit is generally not the role or core competency of the treasury staff. Capital is usually entrusted to them for conservation, not speculation, but pressure on treasury operations to generate profit has created trouble in some companies. In a time of relative stability of interest rates and currency values, a hedge strategy can provide seemingly low-risk profits. When rates turn wrong rapidly—particularly when leverage is employed to multiply the effects of the strategy—things can go wrong very quickly. If panicking employees try to restore losses with "double or nothing" trades, the results can become catastrophic.

### Coverage Considerations

Even in organizations with well established internal controls, seemingly effective plans can go astray. Faced with the reality of a huge loss and large potential liability, can a company find any help from the standard panoply of insurance coverages?

As with other policies, coverage depends on the circumstances. Assume, despite good controls, that the person responsible for the hedging acted beyond his or her authority by taking on more risk than corporate guidelines allowed. Indemnification for financial loss caused by an employee, in the form of fidelity bonds or crime coverages, will be partially dependent on whether the employee attempted to achieve personal gain from the transactions. Further, the policy definition of this gain often excludes commissions and salary. A "trading" exclusion is a common feature of major crime bonds, although it can sometimes be removed. Merely having the exclusion deleted, however, will not normally solve the problem because the requirement that the employee has sought personal gain may still apply.

If the investment strategy that caused the loss has been employed for the potential benefit of a pension fund, the question of the prudence of the investment will turn on management's potential ERISA liability. Defense expenses related to the action will probably fall under the terms of the fiduciary liability policy.

When one is in the business of providing services in the hedging world, as in a bank or broker-dealer environment, the question of professional liability arises. Investment banks and broker-dealers involved with derivatives are very aware of the distinction between advising and dealing. Many of them say they are acting as dealers, not advisors, and should not be held to a fiduciary's legal standards and duties. This issue has been very cloudy, but the agreements that have surfaced publicly seem to imply a higher fiduciary duty. In this instance, a professional liability policy might respond if a suit is brought against the party allegedly giving faulty advice.

The failure of company management to properly supervise an errant employee trader is probably a wrongful act under most D&O policies, but someone has to make such an allegation against a director or officer to trigger coverage.

In all of these instances, the best protection rests in good controls and supervision from knowledgeable managers and outside advisors. Failing that, one can follow the lead of others and sue the advisor who structured the deal for you, alleging he or she did so without full disclosure of the risks. This is unlikely to succeed in a realm where dealers are focused on clear communications and providing written documentation of their advice.

Does a risk manager need to understand all of the instruments used within his compa-

ny? The simple answer seems to be yes. Underwriters are eager to understand a company's management style, and if the risk manager wants to be able to represent the firm directly to the underwriting community, he or she should be able to define the parameters of the company's derivatives exposures. This probably means being able to describe the major hedging strategies and explain the displays in the financial statements. Coming to grips with this exposure will likely require interviews with in-house treasury or trading management and maybe the bankers who designed the firm's derivatives strategy.

Other than discussions with underwriters, there is no clear consensus about the risk manager's role regarding investment risk. Almost all risk managers have a responsibility for reviewing risk control options in place across their organizations. Most brokers and consultants who traditionally operate in the insurance-related risk management arena do not currently provide products or services that deal directly with derivatives risk. Risk managers and their broker-advisors, however, need to understand investment risks and be comfortable with the internal controls because their underwriters are going to ask questions. Insurance carriers will be particularly interested in management oversight of derivatives risk and strategies employed to minimize it. The broker/client team should understand the risks in the company's trading operations to communicate more effectively with underwriters. Another important step is educating management about the limitations of coverages such as employee dishonesty, D&O and professional liability insurance in protecting the organization properly.

Although it is not likely that most risk managers will become actively involved with a company's specific derivatives trades, it is nevertheless important to have sufficient knowledge about these activities to prevent serious consequences and negative headlines. In today's environment, understanding enough about derivatives instruments to ask the appropriate questions and to help establish the company's internal controls and trading limits is critical. ■

12



# Finite Risk Contracts

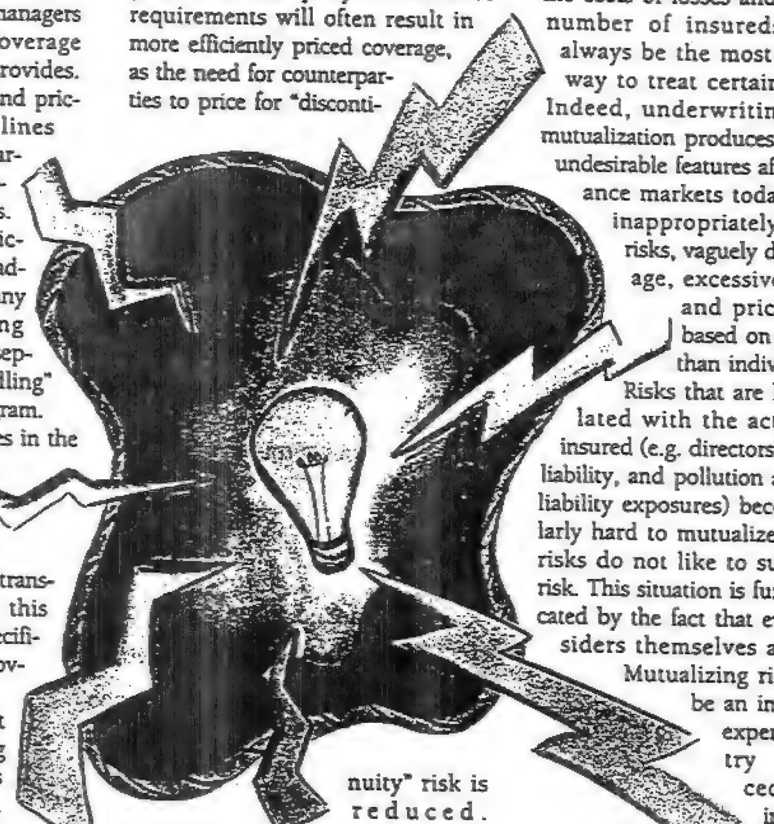
## An Enlightened Approach

by Michael D. Hamer and Thomas R. Dickson

Today, leading-edge insurance programs are shaped by the many imbalances that exist between the protection risk managers need and the coverage that traditional insurance provides. Volatility in the availability and pricing of certain insurance lines reflects the market losses of carriers, rather than the performance of individual insureds. Furthermore, coverage restrictions have been increasing, leading to an "unbundling" of many coverages—that is, writing insurance for each exposure separately, as opposed to "bundling" exposures under a single program.

These manifest inefficiencies in the insurance distribution system are prompting insureds to seek other alternatives. The rapid growth of the alternative risk transfer market bears witness to this search for more effective—specifically long-term—insurance coverage.

Insureds in today's market are increasingly considering their insurance requirements over a multiyear time frame. The rationale for this is that, because it helps companies bear risk, insurance is a form of capital, and it makes no economic sense to arrange annually all of a corporation's capital needs. Most companies' business plans extend over many years; thus, their risk management programs should as well. Furthermore, creating multiyear arrangements for at least part of a company's insurance requirements will often result in more efficiently priced coverage, as the need for counterparties to price for "discontinuity" risk is reduced. (Here "discontinuity" risk refers to the possibility that one party can walk away, leaving the other in a deficit position). Also, the costs of arranging coverage annually are reduced.



BROOKE SCUDDER

The desire to develop more effective insurance coverages also stems from the recognition that traditional "mutualization" of risk—spreading of the costs of losses and risks over a number of insureds—may not always be the most appropriate way to treat certain exposures. Indeed, underwriting linked to mutualization produces many of the undesirable features affecting insurance markets today, including inappropriately unbundled risks, vaguely defined coverage, excessive exclusions and pricing policies based on group rather than individual results.

Risks that are highly correlated with the actions of the insured (e.g. directors' and officers' liability, and pollution and products liability exposures) become particularly hard to mutualize, since good risks do not like to subsidize bad risk. This situation is further complicated by the fact that everyone considers themselves a good risk. Mutualizing risks can also be an inefficient and expensive way to try to match cedents with insurers.

In multiyear insurance programs, the risks need not be so closely defined, so long as they are reasonably consistent from period to period. This permits bundling of exposures, with the associated benefit of "offset" within the

*Michael D. Hamer is chairman of Centre Reinsurance Representatives Limited in London and Thomas R. Dickson is a senior vice president at Centre Reinsurance (Bermuda) Limited in Hamilton, Bermuda.*

aggregate bundle (that is, good results from one of several exposures in a program offset the bad results of another exposure within the same program); taking the same risks over time also makes it easier to build in individual incentives in the form of additional premiums and profit sharing to better align the interests of both the insured and the insurer over time. It also enables the cost of the program to reflect the insured's own loss history rather than that of its peers, and reduces the costs of putting protection in place.

Over time, multiyear programs have evolved, and the range of risks covered has increased. For example, in the past 18 months, dislocations in the property catastrophe markets have propelled the development of multiyear structures for risks with low frequency and high severity. Insureds, unhappy with the volatility in capacity and rating, have sought contracts that provide contractual coverage at a known price for an extended period after a major event.

The following are several key questions for insureds to consider with respect to multiyear contracts, breadth of coverage and other key issues, together with the authors' conclusions.

**FINITE RISK CONTRACTS ARE LONG-TERM—TYPICALLY THREE TO FIVE YEARS—AND PROVIDE CLIENTS WITH CONSISTENT COVERAGE AND STABLE TERMS TO PROTECT THEM FROM POTENTIALLY DAMAGING MARKET FLUCTUATIONS.**

### Multiple-Year Contracts

Q: *Why should I tie myself down with a long-term contract when I'm used to buying insurance on an annual basis?*

A: Most insurance contracts are one year long for no other reason than that is how long the earth takes to complete one revolution around the sun. For companies with a one-year business plan, a one-year contract is sufficient. However, the risks most companies face do not fit easily into discrete one-year intervals. Companies that buy annual coverage often find that volatility is not reduced, only shifted from the year of the loss to the year in which their insurers restrict coverage or demand a payback. In contrast, finite risk contracts are long-term—typically three to five years—and provide clients with consistent coverage and stable terms to protect them from potentially damaging market fluctuations.

*Conclusion:* Multiple-year contracts provide clients with better protection and more stable pricing than annual contracts for many types of risk.

### Broad Coverage

Q: *Across the board, don't conventional products provide more protection than finite risk insurance?*

A: No. In fact, insureds have found that traditional insurance solutions often fail to provide full coverage for losses because they contain exclusions that limit or eliminate coverage for unusual, unpredictable events; if this were not the case, insureds today would not make constant demands for coverages that provide difference-in-conditions protection.

In other cases, the coverage fails to pick up both high frequency and high severity loss events, or the benefit of coverage on one line of business is not offset by the premiums paid on another line of business that does not produce a loss. This creates uncertainty for the insured regarding the amount of insurance available in any particular set of circumstances. For companies in this situation, finite risk and other nontraditional products eliminate confusion by plugging the holes in traditional covers with broad-based, no-exclusion terms for most contracts.

*Conclusion:* In many instances, broad, multiyear coverages provide greater certainty of coverage and a more effective way to protect net income and shareholders' equity.

### Limited Liability to Preserve Financial Security

Q: *Why should I purchase a contract with a liability cap when my company could purchase a policy without one?*

A: Unfortunately, in some cases companies have discovered that so-called "unlimited" insurance coverage unexpectedly becomes very limited when the financial condition of their insurers deteriorates. In other words, all coverage is limited by the insurer's ability to meet its contractual obligations. Policies without liability caps expose every insured to the risk that another policyholder will produce losses that threaten its insurer's solvency.

In the vast majority of situations, aggregate caps allow clients to buy only the amount of coverage they need, especially when analyzed over a multiyear period. Most companies do not really need unlimited coverage, nor do they want to pay for it. In fact, many of the authors' discussions with clients begin by defining the amount of coverage needed over the life of the contract. After determining this, we can develop a cost-effective price based on the term and profile of exposure and the requested level of aggregate limit.

*Conclusion:* Aggregate limits protect policyholders as well as shareholders by allowing broader protection to be offered on a more cost-effective basis. ➤

### Profit Sharing to Align Economic Interests

Q: *With the new accounting rules that are now in place in the United States and elsewhere, isn't profit sharing dead?*

A: On the contrary. The new accounting rules passed in various countries have made the treatment of profit sharing clearer by formalizing the accounting for contracts that contain profit-sharing features. Nevertheless, insureds are best served by focusing on economic issues as opposed to accounting formalities. Profit sharing aligns the economic interests of the insured and insurer.

In many traditional insurance relationships, the company "wins" only if it suffers a large loss resulting in a claim to which its insurer must respond. It "loses" if it does not suffer losses and the insurer then keeps most or all of the premium. Either way, the incentive to keep losses low is reduced. With finite risk contracts, for example, both parties benefit in low-loss cases, so both are motivated to keep losses to a minimum.

Q: *Doesn't my traditional insurer already give me a fair deal?*

Some do, but others don't. In theory, traditional insurers should, at the start of a business relationship, commit to set up "banks" for clients with good loss experience. Unfortunately, many companies have found that insurers do not always set up these banks, or if they do, they use them inappropriately.

For example, in some cases, an insurer will use the bank to cover losses arising from business with other companies in the same industry, or pay out losses from other areas of the insurer's operations as dividends to shareholders or bonuses to management. In other cases, insurers withdraw from businesses, taking the bank with them.

While a policyholder's rates may decline because of good loss history, they are likely to rise more quickly if unexpected losses occur. To eliminate this volatility and uncertainty, attentive insurers will set up reserves against amounts owed under profit sharing so that available funds are returned to policyholders.

*Conclusion:* Profit sharing aligns the interests of insured and insurer, and allows coverage to be offered on a more cost-effective basis.

Beyond the structural benefits discussed above, multiyear programs are individually designed to meet clients' needs. In today's ever-more volatile markets, insureds need risk management solutions that more precisely address their constraints and objectives. By assisting clients in taking a long-range planning approach to better manage their exposures, multiyear coverages represent such a solution. RM

13

# CATASTROPHE INSURANCE FUTURES

by Kathleen McCullough

WHARTON REPROGRAPHICS


CATE JOHNSON

In December 1992, the Chicago Board of Trade (CBOT) unveiled the first catastrophe insurance futures. This financial tool was said to have the potential to revolutionize the way insurers deal with catastrophic losses by giving companies new avenues for spreading this exposure, increased capacity and the ability to freeze catastrophic loss ratios. Despite the potential benefits, the results so far have not been as spectacular as the CBOT first anticipated.

Upon initial inspection, there are several factors that must be overcome if catastrophe insurance futures are to become readily used. For example, many insurance companies will not actively participate in catastrophe futures trading because state regulators prevent them from holding sufficient amounts of futures to hedge their risks effectively. Also, many insurers feel that catastrophe insurance futures are not liquid enough to offer any real benefits. Further, many industry participants who regard the futures as another form of reinsurance cede their exposures to traditional reinsurance markets. Potential accounting problems and a preference for more conservative investments have also steered many potential participants away from catastrophe futures. Other companies simply do not understand the futures process well enough.

*Despite their potential value in hedging loss, catastrophe insurance futures issued by the Chicago Board of Trade face several obstacles before they can be widely accepted by the insurance industry.*

There is a simple way to look at catastrophe insurance futures. They are a financial tool, which if purchased properly, can allow an insurer to offset its catastrophic losses by using capital gains from the rise in the futures index. In other words, the insurers create a hedge by purchasing catastrophe insurance futures. If an unexpected catastrophe occurs, the value of the catastrophe index will rise, and a rise in the price of the

*Kathleen McCullough is 1995 graduate of St. Mary's University in San Antonio. This article is adapted from her paper, "Catastrophe Insurance Futures," which received a 1995 RIMS Research Award.*

Table 1

## PRICING A CONTRACT

To determine the price of a $25,000 catastrophe insurance futures contract, the CBOT uses the following formula:

$25,000 $\times$ (Incurred Catastrophic Losses/Estimated Property Premium)

In this example, assume that the companies in the index pool will experience $300 million in catastrophic losses during a quarter, that 75 percent of those claims will be reported by the end of the quarter and that the companies will receive $4 billion in premiums. With these figures the contract price can be calculated:

1) $25,000 $\times$ (($300 million $\times$ 75 percent)/$4 billion)
2) $25,000 $\times$ ($225 million/$4 billion)
3) $25,000 $\times$ .05625 = $1,406.25 per contract

With the contract price, an insurance company can determine how many contracts it needs to hedge its catastrophic losses with this formula:

Number of Contracts = (Earned Premium/Contract Size) $\times$ (Hedged Losses/Percent Reported)

For example, a company with $15 million in earned premium value using $25,000 contracts could hedge 100 percent of its potential catastrophic losses. If we assume 75 percent reported losses, we get the following figures:

1) ($15 million/$25,000) $\times$ (100 percent/75 percent)
2) $600 $\times$ 1.33333 = 800 contracts needed

futures will help offset the unexpected cost of the losses. Along with the investment risk, another danger to insurers is that the hedge will only cover losses to the extent that the insurer's loss experience matches the loss experience of the companies that comprise the index.

In addition to investors willing to take a chance on these futures, the CBOT has identified four primary groups that represent potential markets for catastrophe insurance futures: large insurers, reinsurers, small insurers and building supply firms and construction companies.

Quarterly futures contracts and options are available for the Eastern, Midwestern and Western regions of the United States. Investors can also purchase an annual Western regional contract and quarterly national futures for all 50 states. The contracts cover aggregate losses for the period in the following lines of insurance: private and commercial automobile damage; fire; multiple peril for homeowners, commercial buildings and farmowners; earthquake; and commercial inland marine.

### CREATING AN INDEX

Catastrophe insurance futures are based on an index that measures the level of catastrophic loss in the property/casualty industry. Due to the fact that this index is the basis for the value of the futures, it is essential that the index accurately and predictably represents the catastrophic losses of the property/casualty industry.

The CBOT uses ISO DATA, a unit of the Insurance Services Office Inc., as its source of statistical information. ISO DATA gathers, stores and calculates statistical and actuarial information for regulators and insurers. One of the primary reasons for using ISO DATA is that ISO is the largest insurance advisory organization in the country, and it is regarded for accuracy and integrity.

Before the contract is listed on the CBOT, at least ten of the 100 insurance companies that report to ISO DATA are chosen to comprise the index. These insurers are selected based on the size of their business and the diversity of the risk they represent. ISO DATA also weighs loss information for the various states and lines of insurance. The information used to settle a particular contract is based on the losses that occurred in the previous quarter. These losses are reported to ISO DATA by the end of the contract month. This information is then statistically weighted and plugged into a formula to determine the index value. Table 1 demonstrates how the contracts are priced.

Due to the fact that catastrophe insurance futures are traded on the open market, the price at any given time should reflect the ultimate settlement price. Thus, if the insurers in the pool experience higher than expected catastrophic losses during a quarter, the index should rise proportionally. For example, if there is a catastrophic loss that causes catastrophe claims in the pool to exceed expected claims by 45 percent, then it is likely that the catastrophe insurance futures price will rise approximately 45 percent. If properly hedged, the insurers and reinsurers who invested in this market should have a capital gain to pay the unexpected rise in claims (see Table 2). If the actual losses paid are less than the expected losses for the pool, the price of the catastrophe insurance futures will probably drop.

The first step in utilizing the futures market is to discern which contract has the underlying risks most closely related to the insurer's risk portfolio. For example, if the insurer's book of business is heavily concentrated in the eastern part of the country, the insurer may wish to purchase Eastern Catastrophe Insurance futures. If the risk is diversified over the entire country, the insurer may wish to purchase the National Catastrophe Insurance futures.

### POTENTIAL PLAYERS

Large insurers have very large premium volumes each year and many of them carry a large amount of catastrophic exposure. To guarantee that they remain profitable and solvent,

Table 2

## HOW LOSSES AFFECT THE INDEX

If a catastrophe produces larger than expected losses, the value of the index will increase according to the following formula:

**Contract Value x (Total Reported Losses/Total Estimated Premium)**

Now, use the expected losses, claims percentage reported and premium figures from Table 1 and assume that losses for the index companies were $80 million more than expected; the contract price would follow the formula:

1) $25,000 x (($380 million x 75 percent)/$4 billion)
2) $25,000 x ($285 million)/$4 billion
3) $25,000 x .07125 = $1,781.25 (final settlement price)

Compared to the initial futures contract price of $1,406.25, this final settlement price represents a $375 gain per contract. If an insurer purchased 800 contracts, it would realize a total gain of $300,000 from the trade.

these insurers must look for ways to shield themselves from unexpected catastrophic losses. Traditionally, insurers have turned to the reinsurance market for protection. Due to capacity shortages and the high cost of reinsurance, many insurers are looking for alternative ways to hedge this exposure. If properly used, catastrophe insurance futures could provide such a hedge. Because a majority of the large insurers represent a well diversified portfolio of risk, it is likely that their loss experiences will be correlated with the catastrophe index. Large insurers have the historical statistical information necessary to compare their loss history for many years against the catastrophe index.

Similarly, because most reinsurers deal with several insurers, they also represent a diversified set of risks. Since their risks are generally spread over a wide geographical area, reinsurers should experience the same benefits from the catastrophe insurance futures as large insurers. In theory, the catastrophic loss experience of the reinsurance industry should mirror the catastrophic losses within companies' U.S. portfolios. Price movements in the catastrophe insurance futures should offset unexpected losses in the same fashion that they protect primary insurers.

Although small insurers generally do not enjoy the benefits of well diversified geographic risks to the extent that large insurers and reinsurers do, they can still benefit from the use of catastrophe insurance futures. Their purchases will take more research and planning, and they may turn to regional contracts instead of national contracts. Small insurers will also have to look closely at their historical loss runs in relation to the catastrophic index to determine to what level their losses are correlated. Once this correlation is established, the insurer should be able to determine which contracts can be purchased to form an effective hedge for its book of business.

Building supply firms and construction companies could trade futures in an opposite position from the insurers. These companies could potentially invest against the insurance industry because a segment of their profit cycle is inversely related to the insurance industry in the wake of a catastrophe. When a natural disaster strikes, the property insurers generally incur large unexpected losses, and the building suppliers and contractors usually experience a rapid increase in demand for their services to repair damaged buildings. Conversely, if the insurers experience an unexpectedly low level of property claims in a given area, the contractors will experience a drop in demand for repairs to insured facilities. For this reason, building supply firms and construction companies may take investment positions opposite those of the insurance industry. Not only can these be profitable investments for the building suppliers and construction contractors, they can also add depth and liquidity to the catastrophe insurance futures market.

While each group has its own concerns regarding entrance into the futures market, all must be concerned with determining the level of correlation between their own catastrophic loss experience and that of the index. If there is a high level of correlation, then a major barrier for market entrance is removed.

Each of the potential investors stands to gain from effectively investing in the catastrophe insurance futures market. As more and more investors become part of the market, its liquidity and effectiveness will improve.

### OBSTACLES TO SUCCESS

There are currently several major problems with the use of catastrophe insurance futures. These issues include accounting, regulatory and liquidity concerns. Stigmas and psychological concerns about catastrophe insurance futures also play a role in the sparse use of this alternative.

*Accounting concerns:* The fact that gains in catastrophe insurance futures are considered investment income creates a perplexing accounting problem for insurers. They want to show the income as counteracting their underwriting losses, but current accounting principles dictate that futures earnings should be reported as investment income. There are no standards governing accounting procedures for catastrophe insurance futures. This has created a sense of confusion within the industry that has caused many to abandon consideration of the concept.

*Regulatory issues:* The regulatory battle is complex because many regulators have yet to decide if catastrophe insurance futures are a viable means of hedging. In fact, many regulatory

bodies are waiting to see what type of experience insurers are having with the use of the futures before they pass any concrete regulations.

Illinois was the first state to specifically address regulatory standards for catastrophe insurance futures. In other states, catastrophe insurance futures are addressed by investment-basket clauses. These clauses prevent insurers from developing "baskets of investments" that are large enough to allow the insurers' profitability level to become contingent on the outcome of their investments for the year. Regulators have placed limits on the riskiness and amounts of investments insurers can purchase. Because futures are considered risky, many states have set limits on the amount of these investments insurers can hold. This amount is generally significantly lower than the amount needed to construct an effective catastrophe hedge.

While the lack of regulation has been a hindrance for some wishing to enter the market, in other cases it has not been a factor. Many of the reinsurers that have considered catastrophe insurance futures as a viable option are located offshore. Thus, decisions made by state regulators do not affect them. Self-insureds, captive insurance companies and other nontraditional risk management sources can also enter the market without direct permission from regulatory bodies, but these sources are only a small percentage of the potential traders in the market. For the market to grow, regulators will have to provide a favorable scenario that may entice the insurance industry. At the same time, regulators will have to be able to assure the public that catastrophe insurance futures will not jeopardize the financial stability or liquidity of the insurance industry.

*Liquidity shortages:* The CBOT has limited participation in this market by restricting the size of the hedges insurers can construct. The primary reason for this rule is that the market for catastrophe insurance futures is still relatively small. Because the CBOT wants the futures' price to be based on the movements of the catastrophic index, rather than demand in the market, large purchases, sales or holdings are not permitted at this time. This causes a problem for large insurers because they cannot construct an efficient hedge with the current limitations. It also creates a problem for those who would like to buy or sell their holdings on short notice. The liquidity problem will not be solved until more investors begin to participate in the market. Yet, as with many of the issues surrounding these futures, people will not begin to participate until the liquidity problem is solved—creating a type of Catch-22.

*Lack of demand:* While most people in the insurance industry agree that the basic principles for the catastrophe

insurance futures are viable, many do not see how futures can fit into their situation. Others simply are content with the current state of their retention and risk transfer systems.

Similar arguments have been made after the initial introduction of other financial instruments. For example, there was a lack of interest when futures linked to U.S. Treasury bond interest rates were introduced in the 1970s. For the first few years after they were introduced, trading was very light. At the time, many investors could not rationalize a need for these futures. Now they are the most widely traded contracts on the CBOT. It may be a matter of time before the insurance industry can see the benefits of catastrophe insurance futures. If there is a sound real-life example of their successful utilization, more prospective investors will recognize catastrophe insurance futures as a legitimate financial tool.

*Conservative cultures:* For the most part, individuals in the insurance industry do not strive to radically change the status quo. They tend to move with caution, taking time to be sure that their decisions will provide the best long-term benefits possible. This has generally been considered an asset due to the fact that it prevents insurers from making rapid decisions that could jeopardize liquidity. Because catastrophe insurance futures have not been a part of the traditional way of doing business, insurers may be reluctant to bring futures into their business plans. Catastrophe futures simply do not appear to fit their corporate cultures or philosophies. Many companies pride themselves on the simplicity and strength of their current condition and are not willing to jeopardize their financial stability with a new "trendy" investment technique that has not proven to be successful over the long term.

*Lack of education:* One of the problems facing the CBOT is that many potential investors either have little knowledge about catastrophe insurance futures, or that investors have misconceptions about the futures. Educating the insurance community could help expand the market for catastrophe insurance futures and could also serve as a means for the CBOT to refine the product. There are many issues that still must be ironed out to make the process run smoothly. By interacting with potential investors, the CBOT could help them develop viable solutions for these problems. This process could also help the CBOT develop other insurance futures.

Many people feel that in time, the catastrophe insurance futures market will grow. Markets and needs are continually evolving. As its capacity is stretched to new limits, the insurance industry will have to develop innovative new tools to fund these risks. Financial alternatives such as catastrophe insurance futures have the potential to fill such a gap. ■

14



DRAFT

# CORPORATE INSURANCE AND RISK MANAGEMENT: A STUDY OF BRITISH PETROLEUM

Neil A. Doherty
and
Clifford W. Smith Jnr.

The authors are respectively from the University of Pennsylvania and the Univeristy of Rochester. We wish to thank Judith Hanratty and Rodney Insall of British Petroleum for many discussions and insightful comments.

## INTRODUCTION

Insurable events such as product liability suits, toxic torts, physical damage to corporate assets, etc. represent a major production cost to U.S. industry. Some years ago, Mayers and Smith (   ) noted that corporate insurance costs alone were of a comparable order of magnitude to dividend payments. Yet this comparison understates their current importance. With the continued expansion of the tort system (Huber & other cites) and the allocation of costs of cleanup of sites containing hazardous wastes imposed under "Superfund" (cite), these costs have risen substantially in the past decade. Moreover, many corporations self insure a major portion of such costs, and these costs are not always apparent ex ante.

The traditional corporate strategy has been to insure these events. According to the insurance literature, this strategy is based on a somewhat generalized notion that risk is costly to the firm and that insurance provides an appropriate hedge.[1] However, it is normal for firms to retain part of their loss exposure by means of a deductible or by purchasing insurance in the excess risk market; thus only larger losses are insured. This widespread convention is rationalized in the insurance literature[2] and is based on loss distribution characteristics that are typical in many corporate loss distributions. For medium to large size firms, small losses (small fires, employee injuries, vehicle crashes, etc.) occur with regularity and their total cost is predictable within tolerable limits. Retaining such losses introduces little uncertainty in the value of the firm and such losses can be left uninsured. This strategy even has a name, "trading dollars".[3] Larger losses are a different kettle of fish. These tend occur with much lower frequency and can have an appreciable impact on the value of the firm. On the notion that risk is costly, such losses should be hedged by insurance if the transaction costs are not too high.

---

[1]See for example, Williams and Heins, 6th ed, part C; Rejda, 3rd ed, Chpt.3.

[2]E.g. Rejda, 3rd ed P.52-53; Williams, 6th ed. Part C, especially Chpt 13.

[3]Since small losses are predictable in aggregate then insurance would simply involve a trade of known dollars of premium for (relatively) known dollars of loss settlements.

Recent finance literature has provided some rationale for why risk is costly to a firm's stakeholders (Mayers and Smith, Main, Shapiro and Titman, Smith and Stultz, etc.). Risk related costs arise from financial distress, convexity in tax functions, and agency costs, etc. Moreover, the cost of bearing risk to the various stakeholders interacts with other financial characteristics, notably capital structure. These considerations suggest that the optimal risk management and insurance strategy, cannot be based on a generalized notion of corporate risk aversion but must address the effect of risk on the value of the various stakeholder claims and how such effects filter through to residual claimants.

There are other economic factors which have received less attention. Insurance markets are subject to various imperfections. Rather than be considered as a single market, the insurance market is better conceived as a network of interlocking markets where cells are defined by line of insurance, size of exposure, location, seniority of insurance claim, etc. The degree of competition, and thereby prices, varies across this network. Further imperfections are present. Insurance contracts are subject to enforcement losses and costs. All insurers have non zero probabilities of insolvency and, even absent financial distress, potential disputes between the parties reduce the net expected payoff under the insurance contract. Such enforcement costs can become important when prospective losses are very large. Insurance markets have limited capacity to write large exposures and many potential corporate losses are so large that insurance is simply unavailable. Finally, insurance usually is bundled with ancillary products which include risk assessment and monitoring and loss settlement services.

In this paper, we address these economic issues to the insurance and risk management strategy of a major corporation. The choice of subject, British Petroleum (BP), is of particular interest; it is a very large corporation with capitalization rivaling that of the section of the insurance market from which it has purchased insurance. It is a relatively undiversified firm in its product mix, it is predominantly an oil company with both upstream and downstream activities.[4] Its main businesses exploit a limited range of complementary technologies and the nature of its organizational capital is

[4] Most unrelated businesses were spun off or sold during the 1980's.

readily understood. The nature and size of its activities generate possible losses which range from small routine losses (collisions of road tankers) to those which are so large (environmental damage from oilspills) they stretch or exceed the capacity of the insurance market to insure them. Thus, the case provides an opportunity to test the limits of the insurance and for comparing the financial and organizational capital of the policyholder with that of insurers in order to derive comparative advantage in risk bearing. While the risk management strategy derived for BP is peculiar to its corporate and market circumstance, it does nevertheless serve to illustrate the fragility of conventional wisdom. The interaction of economic factors leads to a strategy which turns conventional wisdom on its head. Rather than retain small losses and insure the larger ones, BP has done exactly the reverse.

## II. THE BENEFITS AND COSTS OF CORPORATE INSURANCE

### II.a. THE BENEFITS OF CORPORATE INSURANCE

The nature of insurance is that, by the pooling of many exposures, relative risk is reduced. This benefit is reaped in mutual insurance companies simply by combining many policies. For stock firms, the sub-division and trading of residual claims on the insurance pool in capital markets, provides further opportunity for diversification. Thus, insurance may be thought of the provision of low cost diversification services. This is of value for those who hold undiversified asset portfolios and who cannot issue low cost claims against these portfolios. Individuals and small privately held firms typically fall into this category. Publicly traded corporations differ in that their ownership claims are subdivided and traded in capital markets. Public ownership provides a close substitute for insurance. Consequently, insurance will have little or no effect on the firm's cost of capital. Insurable losses may represent risk which is diversifiable in the capital market and therefore commands no risk premium. In this case, the discount rate is unaffected by the purchase of insurance. Alternatively, insurable losses may represent systematic risk and command a risk premium in the firm's cost of capital. In this case, insurance will affect the firm's cost of capital, but the insurance premium paid to the insurer

will include an appropriate risk premium, thus neutralizing any possible gain from competitively priced insurance. (see Doherty and Tinic 1982, Main 1982, Mayers and Smith 1982).

The foregoing suggests that any gain in value from the corporate insurance for widely held firms, must arise from an increase in expected earnings. In this section we will summarize these effects.

II.a(i) Insurance and the Costs of Financial Distress

Typical insurable events include the physical destruction of assets (plant, buildings, vehicles, etc), loss of income that would have been generated by those assets had they survived, and liability awards made against the policyholder. Since senior claims on the firm are not excused in the event of uninsured losses, the losses fall deadweight on residual claimants. On realization of an uninsured loss, the value of equity will fall and leverage will increase. If these events are insured, the loss in value of equity is hedged. With a perfect hedge (i.e. full insurance), equity will be restored to its preloss value (ceteris paribus) and consequently the preloss capital structure will be restored.

But this assumes that the firm does not become insolvent. With insolvency, uninsured loss will be put to senior claimants, at least in part. Consequently, if an insurance policy is in place, the probability of insolvency is reduced since the senior claimants effectively have prior claims on the proceeds of the insurance policy. From this observation, two benefits to corporate insurance can be identified (cite Mayers and Smith, etc.). The first is straightforward. By reducing the probability of insolvency, insurance reduces the expected costs of bankruptcy and financial distress. Since these costs are borne ex post by senior claimants, then ex ante the price of senior claims will rise if corporate assets are insured.

The second benefit from insurance arises from the joint effect of risk, limited liability and leverage on project selection (cite Myers, JFE 1977?). Having raised debt, the firm's owners have an incentive to select high risk projects since they will exclude from their project criteria the value of the default put which is held by debtholders. By excluding the default put and thereby distorting the

project selection criteria, the firm might select negative NPV projects or reject positive NPV projects. Such behaviour is anticipated by rational bondholders and will be reflected ex ante in the bond price. Thus the loss of value from distortions in project selection falls ex ante on shareholders. Insurance can be used to resolve such time inconsistent incentives. First, insurance reduces the risk of insolvency for any given portfolio of projects selected by the firm.[5] This will reduce the value of the default put and will thereby reduce distortions to project selection. Second, there is a class of prospective project selection decisions that is triggered by the insured event. These decisions relate to the replacement of corporate assets that have been destroyed (e.g. rebuilding the destroyed factory). As shown by Mayers and Smith (JRI article), insurance can be used to bond the shareholders to select all (and only) positive NPV projects.

II.a(ii). Insurance Facilitates writing of Managerial Compensation Contract and Contracts with other Claimholders.

Writing effective managerial compensation contracts, as with other principal/agent contracts, requires an appropriate balance of risk sharing and efficiency. The usual configuration is that shareholders have a comparative advantage in risk bearing over managers since the latter have a heavy investment of human capital in the firm and cannot (at low cost) issue claims on that capital. On the other hand, efficiency considerations require that compensation be linked to appropriate performance measures such as the share price or other measures that link agent performance with the welfare of the principal. The problem is that performance proxies tend to display considerable variation that is unrelated to agent performance, i.e. noise. The higher the signal to noise ratio, the more effective the performance measure in motivating appropriate behaviour. One source of noise is uninsured losses. In the absence of moral hazard, insurance will increase the signal to noise ratio and will thereby aid the writing of managerial compensation packages.

[5]Except perhaps if the insurable events display a high negative correlation with other corporate earnings.

This benefit may be reaped at various levels within an organization. For example, senior management compensation often is linked to share value and insurance can be purchased to reduce the noise in the share price. The profit center concept can be used both laterally and vertically within the firm to provide appropriate compensation to divisional management. Correspondingly, insurance may be used to reduce noise in divisional profits and thereby aid the writing of incentive compatible compensation packages with divisional managers. Of course, the firm could use a set of internal transfers to those divisions that suffer loss in order to purge divisional profits of the effect of the loss. However, insurance may have other advantages. Prospective losses are a cost of production. Insofar as local insurance markets are competitive, premiums provide an appropriate ex ante measure of that cost.

Some similar issues arise in writing contracts with other stakeholders. Employees, trade creditors and those holding performance warranties against the firm are affected by the riskiness of the firm's operations. Risk bearing is often costly to such stakeholders (e.g. employees usually have a significant firm specific investment with their employer) and they will extract an appropriate risk premium. This risk premium can be reduced by insurance.

II.a(iii) Service Efficiencies

Besides hedging loss exposures, insurers provide a set of ancillary services. These additional services include actuarial assessment of the loss exposure (estimation of the parameters of the loss distribution), safety and loss control audits, claim handling services, and loss mitigation services. These services usually are bundled with insurance though occasionally they are priced and sold separately. At issue is whether the insurer has a comparative advantage over the customer in the provision of such services.

Insurance markets function efficiently when insurers accumulate large numbers of exposures.

These conditions not only permit effective risk reduction through pooling (when exposures have low correlation), but also permit insurers to accumulate specialized organizational capital. Such portfolios provide data bases that are amenable to actuarial analysis. Thus insurers are able to classify individual exposures and infer loss distribution parameters. From the insurer's perspective, this is necessary to set competitive insurance premiums. Failure by one insurer to categorize individuals according to loss expectancy will lead to adverse selection by policyholders and this will convey a competitive advantage to rivals (Hoy, Crocker and Snow). Thus competition will lead insurers to maintain and improve their data bases and refine their actuarial models.

One might suppose that the insurers' data bases and actuarial assets would convey a comparative advantage over potential policyholders in estimating the latters' loss distributions. For exposures such fires, vehicle collisions, workers' industrial injuries, etc, this will normally be the case. The frequency of such events makes them amenable to statistical analysis. For other exposures, in which new or specialized technology is used, or if losses are infrequent, statistical analysis is less useful and other estimation techniques are required such as engineering risk assessment. Examples include the failure of new pollution remediation techniques, the destruction or failure of satellites from either satellite system faults or destruction of the launch vehicle, injurious side effects from a new class of pharmaceuticals, etc. The firms which acquire the organizational capital required for the develop such technologies will have that same capital available for risk assessment. Insurers rarely will replicate such capital. Thus production firms rather than insurers will have the comparative advantage in risk estimation for such exposures.

Similar issues arise in connection with safety, loss control and loss mitigation. Insurers data bases permit them to identify common types of accidents such as fires, vehicle collisions, workplace injuries, etc and can often advise clients on potential sources of risk and how expected losses might be reduced. Some specialist insurers (e.g. the Factory Mutual Insurance Firms) focus as much on the safety audit as on the provision of insurance services. However the more complex and specialized the technology, the more likely it is that the producer, not the insurer, will have the comparative

advantage in designing and implementing a loss reduction program. For example, an aviation insurer might provide valuable advice to a small airline on safe operating procedures, but is unlikely to contribute significantly to Boeing's or McDonnell Douglas's design plans for a new generation airplane.

One particular area in which insurers have a significant capital investment is the defense of lawsuits. Access to the insurer's lawyers and other defense resources, can reduce the expected costs of third party liability claims. Insurers are constantly defending such cases, whereas individual policyholders will see such cases infrequently. Moreover, the adverse reputation effects which might arise from vigorous defense of a product liability or workers' compensation lawsuit are partly externalized to the insurer who effectively conducts the case. This is particularly useful where third party claimants have a continuing relationship with the firm such as workers' compensation or product liability claims. Reputation is less important when the third party claimant has no contractual relationship with the firm, such as a third party injured in a vehicle collision. The value of the insurer's defense capability will be lower to large or specialized firms that employ an "in house" legal defense capability.

Another ancillary benefit of insurance is that it is sometimes required under financial responsibility laws. For example, in many jurisdictions an operator of motor vehicles must have insurance or show other evidence of financial capability to discharge third party liability claims. Similar financial responsibility requirements are often found in workers compensation insurance. A further example in U.S. is the financial responsibility requirement attaching to the handling of hazardous materials under the Resource Control and Recovery Act of 19.. In most cases the requirements are modest relative to the resources of medium to large size firms. For example, in U.S. liability insurance the requirement is typically in the range of $30,000 to $50,000. Under RCRA, evidence of financial responsibility is required up to *****(check). Thus, this type of spin off from insurance attaches only with the insurance of relatively small losses.

II.a(iv) Tax Benefits.[6]

Many tax jurisdictions permit the deduction of either insurance premiums or of uninsured losses. Given (a) a constant marginal tax rate and (b) actuarially fair premiums, such provisions would have a neutral effect on the decision to purchase insurance; the expected value of tax deduction would be the same with insurance as without. However, both assumptions are strong. With premium loadings (premiums exceed actuarial value) the value of the tax deduction will increase with insurance purchase.

In many jurisdictions tax functions are convex in corporate earnings; ie. they are progressive. Progressivity can arise from increasing marginal tax rates, from the inclusion of tax shields such as depreciation allowances and carry backs, from the absence of interest accumulation of carry forwards, from the Alternative Minimum Tax provisions, etc. Given convexity, a straightforward application of Jensen's inequality confirms that the expected value of taxes can be reduced by reducing the variance of earnings. Given the tax treatment of insurance premiums and uninsured losses, and absent a strong positive correlation between losses and earnings, insurance may be used to reduce earnings and thereby reduce the value of the tax put. This benefit has local characteristics. For a multinational firm, tax may attach in various jurisdictions. The value of the tax put in each jurisdiction will be determined inter alia by the variance of earnings in that jurisdiction. Thus local insurance protection may reduce local tax obligations, even though such policies have little impact on the riskiness of aggregate corporate income.

II.a(v) Insurance as a Source of Funding.

Insurable losses present the firm with a reinvestment decision; should destroyed assets be replaced?. Given the resolution of incentive conflicts described in II.a. above, there will be a need for funds for reinvestment if the NPV is positive. Clearly, insurance competes with debt and equity as a source of financing. In the absence of insurance, the firm will rely on internal or external funds

[6]Many of the tax issues are discussed in Main 1982, and Mayers and Smith ....

and, given the higher transaction costs associated with the latter, there normally will be preference for using retained earnings. If internal sources are sufficient or if the capital costs can be spread over a sufficient time, it may be preferable to tolerate a temporary increase in leverage to avoid the high transaction of a new issue. But if the loss is large and capital costs must be incurred quickly, a new issue will be required. For such losses, insurance provides a leverage neutral source of financing that permits the firm to avoid the transaction costs of hurried new issues following a loss.

Insurance has its own transaction which, in some respects are similar to those of new issues; for example, both insurance and new issues face costs of information asymmetry. Moreover, insurance may have somewhat closer substitutes than new equity or bond issues. For example, some firms hold dedicated lines of credit that may be drawn in the event predefined property or casualty loss. Insurance will be at a comparative advantage as a source of financing for firms with low liquidity and for small firms for whom the transaction costs of new issues are disproportionately high. (cite ?).

## II.b. THE COSTS OF CORPORATE INSURANCE

### II.b(i) Premium Loadings and Transaction Costs.

The insurance premium payable will reflect the insurer's estimate of the expected loss, various administrative cost, the costs of providing ancillary services, a premium required by the insurer's owners for bearing any systematic risk, efficiency losses (moral hazard and adverse selection) and any rent that might be captured by the insurer. The loading in the premium is the difference between the expected loss and the actual premium. Clearly, the higher the policyholder's perception of the loading, the less attractive is insurance. The perceived loading depends on the policyholder's expectation of loss and will reflect rent and transaction costs implicitly as the difference between this expectation and the actual premium.

Some elements of the loading may not be viewed as a deadweight loss to the policyholder. We have already discussed the various ancillary services offered by the insurer; some of the loading might

simply be viewed as the price of these bundled services. However, other aspects of the loading will deter insurance purchase. This is true of insurer rent, administrative costs and the costs of moral hazard and adverse selection. This permits us to offer conjectures on the comparative advantage of insurance compared with other risk management strategies. Moreover, these conjectures add further fuel to the argument that the costs of insurance will rise with the size of the loss insured.

First, the ability of the insurer to capture rent depends on the degree of competition. The insurance market place might be thought of as a network of separate markets delineated by line of business, size of loss, type of policyholder and by location. Some sub markets are highly competitive since they accommodate many insurers selling large numbers of small, homogenous policies which have low correlation and for which expected losses are easily estimated (cite Joskow). For such markets, barriers to entry are low. Since aggregate losses are fairly predictable, modest capital is required to maintain a low ruin probability. Moreover, the routine nature of the business and the presence of established independent selling networks (independent agents and brokers) requires only a modest investment in underwriters, actuaries, engineers, assessors and other technical skills. Such markets include many personal insurance lines, various insurance of small business and also the routine small property and liability losses incurred by large firms.[7]

At the other end of this spectrum, the market for large losses, and for certain special types of risks, has significant entry barriers and supply is distinctly limited. For example, writing pollution insurance requires a considerable investment in a capability to provide environmental audits and risk assessment; there are only a handful of U.S. insurers active in this market (e.g. A.I.G. and Reliance). Very few insurers have been willing to make a market in very large exposures. To do so requires that the insurer invest heavily in establishing reinsurance facilities (this is largely a reputational investment) and to commit a relatively large surplus to maintain an acceptable ruin probability in a

[7]For example, the fire insurer insuring 5,000 independently owned retail grocery stores will significantly improve its diversification by insuring one further policyholder which is a chain owning 4,000 stores. It might further improve its diversification by insuring the smaller losses on much bigger risks, e.g. small fire losses (e.g. not exceeding $500,000) on department stores, electronics manufacturers, etc.

business with a high coefficient of variation. A significant portion of the world capacity for insuring very large and unusual exposures is provided by Lloyds of London. This is, in part, due to its particular organization form. With unlimited liability, Lloyds syndicates can rely on implicit surplus in the form of a call on the wealth of their members. However, since the value of this call depends on the net worth of the members, the membership of Lloyds', and therefore the available supply of insurance, is highly restricted. Thus, the market place for pollution insurance, unusual lines of insurance, high risk lines and for very high levels of insurance of all lines of insurance is characterized by entry barriers and the potential for insurer rents.

II.b(ii) Moral Hazard and Adverse Selection

Moral hazard and adverse selection arise naturally from insurance contracts. Moral hazard will tend to deter insurance purchase for all potential policyholders since the costs of hidden actions will be anticipated in the insurance premium. Adverse selection will tend to deter insurance purchase selectively for those policyholders who are of low risk relative to their class. One control for such problems lies in the rationing of insurance (see for example Stiglitz 1984 ? Rothschild and Stiglitz, 1976, Wilson 1977) and though second best equilibria may be attainable with such devices (Crocker and Snow 1986) there is a sense of throwing the baby out with the bath water.[8] An alternative control device is experience rating under which the premiums for future insurance contracts depend on prior claim experience. This not only encourages safety on the part of the policyholder (Dionne, Rubinstein and Yaari) but also can be used to reveal hidden information on the loss expectancies of different policyholders (Cooper and Hayes, Dionne and Doherty). If losses are frequent, experience rating is effective in motivating care on the part of the policyholder. Similarly, when losses are frequent, then experience will reveal hidden loss characteristics quickly, information asymmetry will be rapidly reduced and the cost of adverse selection will be correspondingly low. Conversely, experience rating is a less effective control when losses are infrequent. Given a robust tendency for loss frequency to

---

[8]In order to make efficient insurance markets one must partially restrict insurance supply.

be negatively correlated with size, it follows that the costs of adverse selection and moral hazard will be higher ceteris paribus, the larger the size of the loss insured.

II.b(iii) The Limits of Insurability and Costs of Enforcement.

The amount of insurance protection that can be purchased for a single event such as a class action lawsuit, loss of an oilrig, etc, is not limited by the financial capability of individual insurers. Insurance is commonly syndicated amongst several insurers or sometimes the policyholder will stack up layers of insurance protection by purchasing separate policies from several insurers (surplus lines insurers). Moreover, the capacity of individual insurers can be extended by selling secondary claims on its policies (or on its whole portfolio) in the reinsurance market. But, even given such possibilities for combining the capacity of several insurers, there are definite limits to the amount of insurance protection that is available in the marketplace for single events or for accumulations of losses to a single policyholder. As a rough rule of thumb, it is difficult to place insurance for losses in excess of $500 million and rare to find protection in excess of $1 billion. Thus, as a practical proposition, insurance is unavailable for losses exceeding $0.5b or so.

Within the absolute limits of insurability the quality of insurance is not constant. Insurance contracts are not fully enforceable against a bankrupt insurer. Moreover, enforcement of the contract against a solvent insurer can entail significant costs since the insurer can dispute liability and/or delay settlement. These costs include legal costs of bringing suit against the insurer, the time value of delayed settlement, smaller settlement than might arise in the absence of dispute, and increases in the costs of writing future contracts between the parties because trust and reputation have been damaged. Various arguments suggest that the costs of enforcement are likely to rise with the size of the claim. First, and perhaps most obvious, is that a large claim might trigger the bankruptcy. But falling short of bankruptcy, large claims often cause significant financial stress to insurers. This is particularly so in the surplus lines market where the variance of liabilities is usually very high relative to the insurer's surplus (equity). Thus events such as the liability insurance crisis of the mid 1980's and the

earthquake and windstorms of 1989, brought several insurance failures notably amongst excess lines insurers and reinsurers (Source?). It is almost trivial to suggest that the willingness of an insurer to incur the adverse reputation costs of disputing claims will be greatest when the insurer is under financial stress.

Other reasons for associating enforcement costs with the size of claims relate to the relative frequency of large and small claims. The distribution of the frequency of losses by size typically is skewed to the right. Small losses occur routinely but large losses are rare. This has two implications. Small losses tend to be repetitions of similar events, vehicle crashes, workplace injuries, small fires, etc. Given the accumulated experience of the parties in negotiating past claim settlements (and the experience of courts in construing contact wordings in connection with these events) there is limited scope for the parties to disagree on what is covered. Large and unusual losses often present unusual facts and challenges to insurance contract language for which there is little experience. The scope for different interpretations, and therefore the expected enforcement cost, is correspondingly large.

The potential for dispute under large claims can be aggravated by reinsurance. If a primary insurer disputes a claim and loses the ensuing litigation, the primary insurer must settle the claim to its policyholder and the reinsurer is then bound to the primary firm. However, if the primary settles a claim without challenge, but the reinsurer considers that the loss was not covered by the primary contract language, the reinsurer might dispute its settlement to the primary. Consequently, when deciding whether to settle or fight claims, primary firms have to anticipate both the chances of being successful in lititgation with their policyholder and whether they can carry their reinsurer in any settlement. Reinsurance usually attaches on large losses and consequently, this reinforces the positive correlation between enforcement costs and loss size.

The frequency of losses also determines the effectiveness of reputation as a control on opportunistic behaviour. The perceived quality of an insurance contract is determined largely by the reputation of the insurer in settling claims. At the annual round of rebidding to renew insurance contracts, the policyholder has considerable new information on the insurer's performance in settling

frequent small claims. The insurer is unlikely to sacrifice its reputational investment in its clients for modest savings in disputing small claims. For large infrequent claims reputation is not such an effective control. Much depends on how client specific is the reputational capital. The expected saving from disputing a large claim may be large relative to the future profit the firm can extract from that policyholder. Consequently, if the reputational cost can be limited to that policyholder, the insurer might well choose to dispute the claim, delay payment or bargain vigorously on the settlement. However, if this behaviour is observed widely (e.g. if the case is litigated) the reputation costs will not be confined to that client and the reputational effects will be somewhat larger. In general, one would expect that insurer is more likely to behave opportunisticly in the event of large claims. This provides an additional reason why enforcement cost would be positively related to loss size.

II.c Summary

A striking pattern that emerges is that both the costs and benefits of insurance bear a relationship to the size of the loss insured. Some the benefits of insurance[9] arise because insurance can reduce risk and thereby reduce the value of option like claims on the firm. The benefits were mainly "global" in nature[10]; ie. they arose from the effect of insurance on the riskiness of aggregate corporate numbers such as the value of the firm or of its earnings. If insurance is to create value, it must be capable of reducing aggregate corporate risk. However, firms of medium/large size typically exhibit the statistical pattern of many small losses (which are readily predictable in aggregate) and few large losses. Since insurance of the small losses will have little effect on riskiness of corporate value and earnings, it will have little effect on such costs as financial distress. On the other hand, if losses that are large in relation to corporate value are uninsured, costs such as financial distress will

[9]Other benefits, such as the value of ancillary services, may also relate to size as argued but the relationship is more complex than for the costs of financial distress.

[10]The exceptions were the tax claims which depend on the riskiness of local taxable income and the ability of insurance to reduce the riskiness of local profit measures and thereby mitigate the risk return trade off in compensating local managers.

be significant. But the same distribution over small and large claims also suggests that the costs of insurance will likely increase with size of loss because of less effective competition for large losses, less effective controls on moral hazard and adverse selection, etc. This relationship to size was very influential to BP's adopted risk management strategy as we shall now see.

## III. BRITISH PETROLEUM.

### III.a Profile of BP and its Loss Exposure

BP comprises four operating companies. BP Exploration, as its name suggest, is the "upstream" company which is responsible for exploration and development of new oil resources. PB Oil is the "downstream" business and its activities range from refining, distribution and retailing of petroleum and oil products. The remaining companies are BP Chemicals (which concentrates in petrochemical, nitrates, acetyls and nitrates) and BP Nutrition, a relatively small company committed to the animal feed business. The assets of the companies include exploration and extraction licenses, a heavy investment in scientific and technical human capital which is specific to the oil industry (though it may be readily transferrable to other oil firms), production and distribution assets including rigs, pipelines, refineries, ships, road tankers, filling stations etc. Overall, the firm has two major concentrations of value, its production licenses and facilities in the North Sea and in the Alaska North Slope. This value concentration, together with its limited range of activities, imply that it is relatively undiversified. This is certainly true with respect to business risk but it is only partially true with respect to insurable risks. Despite the concentration in production, BP has a wide spread of risk at the distribution end. This may be illustrated by the fact that it has some 13,000 service stations in some 50 countries.

Perhaps the most salient feature of BP is that it is big. It is the biggest company in U.K., the second largest European company and one of the biggest in the world. BP's equity capital is

approximately $35 billion[11] and, with a debt/equity ratio of about 0.4, it has outstanding debt in the order of $15 billion. After tax profit has averaged $1.9 billion over the previous 5 years with a standard deviation of $1.1 billion.

BP's loss exposure ranges from very routine small losses to potential losses in the multi billion dollar range. At the low end of this scale, there are many road vehicle accidents, small shipping accidents, isolated industrial injuries, small fires, equipment failures, etc. Of a larger scale, loss scenarios include refinery fires or explosions, loss of offshore oil rigs, minor environmental damage from oilspills, loss of (nautical) oil-tankers. Large losses could include cleanup costs arising from major oilspills (of Exxon Valdez consequence or larger), tort claim for widespread injuries caused by release of toxic chemicals, liability for defective fuel causing major airline disaster, loss of offshore rig with major loss of life (The Piper Alpha case being illustrative). Perhaps one of the largest foreseeable losses to a oil firm would be the withdrawal of operating licenses as a consequence of political decisions made following environmental damage.

A quantitative picture of BP's loss exposure can be portrayed from its estimated loss distribution[12]. Using actuarial techniques various distributions were fitted around past BP loss data, industry data, hypothetical loss scenarios. Curve fitting permits estimation of the tail of the loss distribution despite the absence (or virtual absence) of very large losses in the past loss record. Distributions such as the compound poisson were found to fit reasonable well. From this technique, the following numerical picture of BP's loss exposure emerges.

---

[11]Information is given at late 1990. This is the time at which BP conducted a study of its risk management strategy.

[12]BP employed independent actuarial consultants to estimate the loss distribution. The actuaries had access to extensive industry and BP loss data. In addition, loss scenarios were constructed for possible large and unusual events that were not in the loss record but which nevertheless were considered to be feasible. It would be rare for an insurer to undertake a study of this intensity to estimate the loss distribution for an individual client.

TABLE 1

| Loss Range | Number per year | Average Loss | Total annual cost | Standard Deviation |
|---|---|---|---|---|
| $0 -$10m | 1845 | $0.3m | $52m | $12m |
| $10m- $500m | 1.7 | $40m | $70m | $98m |
| $500m plus | 0.03 | $1000m | $35m | $233m |
| Whole Distribution | 1847 | $0.66m | $157m | |

Putting these figures alongside appropriate BP numbers and alongside appropriate insurance market indicators is useful. First, as noted the maximum loss that can normally be placed in the insurance market is in the order of $0.5b to $1.0b (which explains the segmentation of the loss distribution above). Noting that losses are tax deductible and assuming a 35% tax rate, a loss of ~~$1.0b~~ $0.5b represents less than ~~2%~~ 1% of the market value of BP equity, it represents ~~34%~~ 17% of average annual after tax earning of the last five years and is about ~~60%~~ 30% of the standard deviation of average earnings. Some further indication of the possible effect of insurance on BP is gained by comparing the risk from major insurable exposures with BP overall risk. Note first that the standard deviation of after tax earnings for BP is in excess of $1b. In comparison the standard deviation from self insuring small losses (less than $10m) is only $12m and the standard deviation from insuring larger losses, but still within the insurance market's capacity, is only $98m. Since insurable risk is modest relative to annual earnings one can conclude that self insuring these losses would add little to the riskiness of the value of the firm.

III.b Previous Insurance Purchases

To the present, BP has insured its property, liability and business interruption exposures to the extent that supply has been forthcoming. Liability insurance and business interruption insurance has been placed largely with external insurers, though some insurance has been purchased from an

oil industry mutual, O.I.L. of which BP is a joint owner. O.I.L. in turn reinsures with unrelated reinsurers. Some of property insurance is purchased directly from insurers. Other property is insured with a captive insurer, however, the captive then reinsures excess liability with external reinsurers. Thus, in total, BP has secured considerable insurance protection from external insurers. Virtually all of this insurance relates to the first two layers in Table 1 and most is in the second layer, ie from $10 to $500m. With one or two exceptions, no insurance has been available above $500m thus BP has always self insured in this range. Over a ten year period, BP has paid $1.15 billion in premiums and has recovered $250m in claims (both in present values). This appears to be about a 360% loading on realized losses. One can explain this loading as (a) a reserve against catastrophe losses, as (b) a payment for ancillary services or as (c) a payment for transaction costs and rent for the insurer. Usually, transaction costs and ancillary services normally account for 10% to 30% of premiums. We now consider whether this loading is well explained as a catastrophe reserve. In other words, could the difference of $900m between premiums and losses be explained simply because the price factored in the prospect of large losses (say up to $500m) but, due to chance, these did not arise over the ten year span? A simply calculation will illustrate that this is ~~quite~~ most unlikely.

Suppose the total premiums paid $1.15b (present values) were competitively priced in the sense that it represent the expected value of losses of $880m plus a typical expense loading of about 20%. Since we are dealing in present values we can approximate this as a sequence of ten annual expected losses each having a present value of $88m. The approximation of constant expected present values allows for growth at a rate equal to the discount rate. Note that even though much of the insurance was in the second tier of the table, we cannot compare this $88m with the most recent estimate of $70m for second tier losses since not all losses in this tier were insured. However, we will use the coefficient of variation from the second tier in the table, (1.4) to estimate the coefficient of variation for losses that were actually insured. This yields an estimate for the standard deviation of annual aggregate losses actually insured of $108.4m. Now if losses are time independent, the standard deviation of the aggregate losses for the ten year period is $\sigma_{10}^2 = [(10)(108.4)^2]^{\frac{1}{2}} = \$342m$. Now the

revealed losses for the period is $250m which is 1.84 standard deviations below the supposed total expected losses of $880m. Using standard normal tables, there is only a .033 chance that revealed losses could be that low.[13] This reasoning contains a strong suggestion that insurers were extracting very high rent to insure these losses. Moreover, as we are discussing the larger insurable losses, this observation is consistent with the conjectures offered above that larger insurable losses will command high loading because of monopolistic rent, etc.

An explanation for the enormous rent extracted by the insurers is found partly in the monopolistic conditions in this market and partly in the high enforcement costs. First, very few insurers are willing to make a market in risks of the size and nature of BP's. Much of the market capacity is provided by some of the Lloyd's syndicates. Some idea of the stress a large loss would cause to these insurers can be gleaned by considering that a single $500b loss represents about 8% of the TOTAL premia; 90% of TOTAL profits and 4% of TOTAL surplus, where TOTAL refers to the combined numbers of all syndicates. Noting that only some of these syndicates made a market for BP, the potential for a large loss to cause stress to the insurer is evident. The problem is aggravated somewhat by the unlimited liability structure of the Lloyd's syndicates. Lloyd's closes its accounts after 3 years and distributes surplus. Thus, much of a syndicate's reserves for long gestation losses, such as liability claims, takes the form of a call on the personal wealth of the underwriting members. In recent years, attempts to recover deficits from members have met legal challenges[14] (e.g. malpractice suits against the underwriter). Given costly access to reserves and surplus (and, as

[13]The estimated value of 0.033 is best considered a upper boundary to the possibility that revealed losses could be as low as $250 assuming that the expected loss is $880. The loss distribution is skewed to the right. For such cases, use of the standard normal table will normally overestimate the probability mass in the left hand tail and underestimate the mass in the right tail. An alternative approximation procedure is to use the normal power approximation. (cite Beard, Personnen and Pentikainen)

[14]In the usual market model, monopolistic rent will attract new insurance capacity. The entry barrier here is the equity capital that is made available to absorb abnormal loss experience. The costs of providing this capital are lower for Lloyd's than for Stock firms because the much of the equity is represented by the call on personal wealth and it is costly for polcyholders to access this surplus. (Cliff, this thought probably needs elaboration or leaving out!)

mentioned earlier, costly access to reinsurance), for large losses one would expect that enforcement costs would be high for the policyholder. This prospect is reinforced by the low frequency of large losses which dulls the usual reputational controls on repeated opportunistic behavior. For example, Table 1 show that the $500m plus loss occurs only about once every 30 years. Consistent with the foregoing, for the only large liability loss that BP has filed with these insurers, settlement with the insurers was only reached several years after settement with the plaintiff. The insurance settlement was only 60% of the liability claim and to secure this, BP had considerable legal costs, and devoted much senior management time to resolving this dispute.

III.c Adopted Risk Management Program.

BP's resolution of its risk management program represented a reversal of the conventional wisdom. Conventional risk management practice is to self insure small losses and insure only the larger losses. In BP's case, losses below $10m are to be the responsibility of local operating units who may seek insurance on local markets, or seek competitive quotations from BP's own captive insurer. Insuring small losses has several advantages. Markets for such losses are highly competitive, insurers have a comparative advantage in offering ancillary services and such insurance is often needed to satisfy local financial responsibility requirements. Delegation of the decision to insure small losses facilitates the writing of incentive compatible contracts for local managers since they have the opportunity to reduce expected costs by appropriate safety programs and low price insurance purchases but insurance also permits removal of noise from performance measures. BP's own captive insurer may compete with external insurers for business at this level. However, given the real resource effects discussed in this paragraph, the captive will not enjoy any hidden subsidy that will permit a bidding advantage over external insurers. Thus, small losses will be subject to the pricing, monitoring and loss control imposed by a competitive external insurance market.

Losses above $10m are not to be insured on external insurance markets except in unusual circumstances (e.g. insurance is required under existing bond indentures or joint venture provisions).

This represents a substantive change in policy for losses in the range $10m to $500m which were previously insured; losses above $500m always had been uninsured. The rationale for this change is (a) the marginal impact of such losses on total corporate risk, (b) the lack of competition in insurance markets of this level, (c) the high costs of enforcement, and (d) the absence of a comparative advantage to insurers in safety services and other services that are ancillary to insurance. The absence of insurance poses adds little risk to aggregate firm numbers. However, random uninsured losses can cause significant noise in local and regional performance measures. Since such losses are infrequent and discrete events, this noise can be purged from performance measures on an ad hoc basis; i.e. losses will not be charged against profit centers (and therefore to managerial compensation) except insofar as the loss reflects poor management (e.g. inadequate safety standards, poor postloss mitigation).

The absence of insurance for larger losses shifts the burden of financing losses to conventional funding sources, debt and equity and can be appraised by conventional criteria. For example investment in the reconstruction of destroyed assets requires that the present value be positive and funding is decided by the usual trade off between internal/external funds and between debt/equity. This has the added advantage that risks that are, in principle, insurable are treated consistently with all other sources of corporate risk.

## III.d Towards an Integrated Risk Management Strategy

Consideration of funding links the discussion of insurable risk to more general issues of corporate risk management. Alongside insurable risk, BP is exposed to risk from oil price fluctuations, currency fluctuations, interest rate changes, demand shifts and the like. For many of these risks, hedging instruments are readily available (oil futures, interest caps, etc.). The optimal hedge on each exposures depends on its correlation with other sources of risk and on the availability of hedges for the other exposures (Mayers and Smith JPE 1983, Doherty and Schlesinger JPE 1983, Brys, Crouhy and Schlesinger 1991). Consequently, the unhedging of insurable risk (cancellation of insurance) has significant implications for BP's overall risk management strategy. We will not address the optimal

design of an integrated risk management strategy here. However, we will conclude by considering some approaches that jointly address insurable and other corporate risks.

The rationale for the cancellation of much of BP' insurance was that, in addition to being too expensive, it was ineffective as a hedge. The ineffectiveness of insurance arose because the range of losses that are within the capacity of the insurance market can be absorbed by BP without significant increase in risk. However, BP is still exposed to the prospect of very large losses (beyond the insurance market capacity) that can cause significant loss of equity value. For example, a major oilspill and the increased regulatory scrutiny that might follow, could cause BP a loss of value comparable to a oil price decrease of several dollars. But unlike the oil price drop, a dedicated hedge is not available to cover the very large oilspill and its consequences. However, other strategies are available.

First, while the simple hedge exhibits perfect negative correlation with the underlying variable, partial hedges can be assembled with less than perfectly correlated variables. The very large losses that can cause significant risk to BP, can impact oil prices. For example, a oilrig fire of the magnitude of the Piper Alpha disaster in the North Sea, can give rise to new safety regulations that increase production costs and thereby affect oil prices. Similarly, a repeat of the Exxon Valdez or a similar event in Alaska, could conceivably lead to withdrawal of operating licenses not only for the implicated producer but for all producers. This would not only affect the market values of oil companies but could also affect oil prices. Such possibilities define potential hedging instruments, one being the equity of rival firms, the other being oil futures. Thus, the same instruments might value in hedging insurable and non insurable risks.

Another way of jointly addressing the effects of insurable and non insurable risk is to reduce leverage. The costs of financial distress and agency costs arise from the interaction between risk and leverage and these costs can be reduced by lowering leverage. Thus, the optimal leverage for a firm will be lower for the self insured firm than for the insured firm. In this sense, leverage provides a substitute for corporate insurance.

A more subtle approach is to change leverage ex post, i.e. following a loss. By issuing convertible bonds the firm can simultaneously change the probability of bankruptcy and induce a change in leverage conditional upon an uninsured loss. The normal convertible bond is the combination of a non convertible bond and a call option on the stock, the option being held by the bondholder. The conversion option introduces upside risk to the bondholder who will exercise the option only when stock values rise. The presence of the option increases the value of the bonds and reduces the fixed income commitment thereby reducing the probability of bankruptcy. An alternative approach is to reverse the conversion option, i.e. the bondholder holds a long position in the bond however the firm holds an option to convert to stock. Clearly, the firm will exercise its option only when prices fall. Thus such instruments can be used to trigger a conditional reduction in leverage when equity loses value (e.g. when oil prices fall or when an "Exxon Valdez" type of event occurs). The reverse conversion option thereby permits the firm the immediately benefit of debt financing but reduces the probability of bankruptcy by conditioning future leverage on the occurrence of loss. In the limit, if all debt is issued on this basis then the probability of bankruptcy would be reduced to zero.[15]

[15](Cliff, no doubt such an instrument could be assembled in other ways)

BP FINANCIAL DATA

15



# Balance Sheets

| | | | | | £ million |
|---|---|---|---|---|---|
| | | Group | | Parent | |
| **At 31 December 1989** | Note | **1989** | 1988 | **1989** | 1988 |
| Fixed assets: | | | | | |
| Intangible assets | 16 | **1,672** | 1,874 | **—** | — |
| Tangible assets | 17 | **19,285** | 18,926 | **—** | — |
| Investments | 18 | **1,497** | 1,437 | **1,755** | 2,333 |
| | | **22,454** | 22,237 | **1,755** | 2,333 |
| Current assets: | | | | | |
| Stocks | 19 | **3,381** | 2,503 | **—** | — |
| Debtors | 20 | **5,361** | 4,243 | **3,854** | 4,357 |
| Investments | 21 | **151** | 157 | **—** | 20 |
| Cash at bank and in hand | | **268** | 183 | **3** | 3 |
| | | **9,161** | 7,086 | **3,857** | 4,380 |
| Creditors – amounts falling due within one year: | | | | | |
| Finance debt | 22 | **2,531** | 2,319 | **—** | — |
| Other creditors | 23 | **7,037** | 5,659 | **811** | 883 |
| Net current assets | | **(407)** | (892) | **3,046** | 3,497 |
| **Total assets less current liabilities** | | **22,047** | 21,345 | **4,801** | 5,830 |
| Creditors – amounts falling due after more than one year: | | | | | |
| Finance debt | 22 | **5,758** | 4,864 | **—** | — |
| Other creditors | 23 | **1,936** | 1,719 | **18** | 8 |
| Provisions for liabilities and charges: | | | | | |
| Deferred taxation | 9 | **451** | 389 | **—** | — |
| Other provisions | 24 | **2,461** | 2,241 | **—** | — |
| **Net assets** | | **11,441** | 12,132 | **4,783** | 5,822 |
| Minority shareholders' interest | | **656** | 614 | **—** | — |
| **BP shareholders' interest** | | **10,785** | 11,518 | **4,783** | 5,822 |

| | | | | | |
|---|---|---|---|---|---|
| **Represented by** | | | | | |
| Capital and reserves: | | | | | |
| Called up share capital | 25 | **1,346** | 1,536 | **1,346** | 1,536 |
| Share premium account | 26 | **1,752** | 1,685 | **1,752** | 1,685 |
| Capital redemption reserve | 27 | **197** | — | **197** | — |
| Reserves | 28 | **7,490** | 8,297 | **1,488** | 2,601 |
| | | **10,785** | 11,518 | **4,783** | 5,82 |

**Peter Walters**, Director.

**Lindsay Alexander**, Director.

15 February 1990

# Group Income Statement

| For the year ended 31 December 1989 | Note | 1989 | 1988 |
|---|---|---|---|
| | | | £ million |
| **Turnover** | 2 | 29,641 | 25,922 |
| Replacement cost of sales | | 22,095 | 19,330 |
| Production taxes | 3 | 1,242 | 566 |
| **Gross profit** | | 6,304 | 6,026 |
| Distribution and administration expenses | 4 | 3,690 | 3,333 |
| Exploration expenditure written off | | 480 | 495 |
| | | 2,134 | 2,198 |
| Other income | 6 | 803 | 693 |
| **Replacement cost operating profit** | | 2,937 | 2,891 |
| Stock holding gain (loss) | | 390 | (232) |
| **Historical cost operating profit** | | 3,327 | 2,659 |
| Interest expense | 7 | 794 | 582 |
| **Profit before taxation** | | 2,533 | 2,077 |
| Taxation | 9 | 744 | 823 |
| **Profit after taxation** | | 1,789 | 1,254 |
| Minority shareholders' interest | | 45 | 44 |
| **Profit before extraordinary items** | | 1,744 | 1,210 |
| Extraordinary items | 10 | 390 | — |
| **Profit for the year** | | 2,134 | 1,210 |
| Distribution to shareholders | 11 | 795 | 823 |
| **Retained profit for the year** | | 1,339 | 387 |
| | | | |
| **Earnings per ordinary share** | 12 | 31.8p | 20.0p |

# Group Reserves

| | Note | 1989 | 1988 |
|---|---|---|---|
| Group reserves at 1 January | | 8,297 | 7,802 |
| Exchange adjustments | | 268 | 117 |
| Shares purchased from KIO | 27 | (2,423) | — |
| Retained profit for the year | | 1,339 | 387 |
| Other movements | | 9 | (9) |
| **Group reserves at 31 December** | 28 | 7,490 | 8,297 |

16

Case 1



# ABC MANUFACTURING COMPANY

## Description of Operations

ABC Manufacturing Company is a manufacturer of fine furniture designed to resemble the furniture used in various historical periods. The company's furniture is sold by approximately fifty dealers throughout the United States and in several foreign countries. Dealers are carefully selected by ABC to be sure they maintain the high quality image that ABC has established in its 100-year history of making fine furniture. Most dealers sell only ABC's furniture, but a few also sell noncompeting lines of similar quality. Dealers who sell only ABC's furniture are franchised to do business under ABC's trade name and would appear to the public to be branch stores rather than independent businesses.

ABC's factory is a three-story building of heavy timber (mill) construction located in a New England town of 2,500 population. The first floor is used for warehousing and office space, and the two upper floors house the production facilities. The building is seventy-five years old and has been occupied by ABC since its construction. The local fire department is above average for towns of this size, but its water supply

system does not have adequate reserve capacity to maintain pressure for fighting a major fire over a period of several hours.

Adjacent to the factory building is a large frame structure used by ABC to store seasoned lumber for its furniture. Lumber is first air dried in the yard adjoining the shed, with the exact drying period depending on the kind of wood involved. The air-dried wood is then kiln dried in ABC's own kiln, after which it is stored in the frame storage building until needed. ABC's power plant is located in a brick building about fifty feet from the main plant. It furnishes heat and steam for all operations. Water for the steam boilers is drawn from a large river adjacent to the power plant. Electricity is purchased from a public utility.

Due to the length of the seasoning period, ABC usually has on the premises a supply of lumber adequate for four months of operations. Much of the lumber is purchased within a 200-mile radius of the factory, but substantial amounts are purchased from more distant sections of the United States or from foreign countries. For example, mahogany lumber, which is used in almost half of ABC's products, is imported by ABC from Latin America and Africa. Most of the lumber arrives at ABC's yard by railroad and is shipped FOB point of origin. The imported lumber is shipped by water to the nearest port, approximately 100 miles from ABC's factory, and then by train to the factory. Some locally produced lumber is delivered to ABC by truck, FOB ABC's yard.

Workers in ABC's factory use some small power tools, such as saws, planers, lathes, and similar equipment. However, the production process is primarily manual. Highly skilled craftsmen build the company's products, using many of the same techniques that were used two centuries ago. Furniture finishing is done by hand, and most of the stains, varnishes and other finishing materials are compounded by ABC's employees according to the company's proprietary formulas. ABC's labor force is very stable because each craftsman undergoes a lengthy apprenticeship in the plant and because ABC pays wages that are relatively high by the standards of the community. Management considers the high wages to be necessary because of the time and expense required to hire and train a replacement for a craftsman who leaves the firm.

Some finished furniture is shipped by railroad, but most of it is shipped by contract carrier trucks. The contract with the trucker does not include any provision relative to liability for damage to goods in transit. Export shipments are transported by truck to the nearest port and by water to the country of destination. All shipments are made FOB purchaser's warehouse.

All sales, including export sales, are made on open account and only to ABC's established dealers. Export accounts usually are denominated in and payable in the currency of the importer's country. Accounts

receivable, on the average, are equal to about one-eighth of annual sales. Accounts receivable records are kept on ABC's computer in the office section of the first floor of the factory building. The computer is also used for inventory and production control, payroll management, and other accounting functions. The computer equipment is owned by ABC. The office is cut off from the warehouse section of the first floor by a wood partition, and the computer room is cut off from the balance of the office by a similar partition. The office and computer room are air conditioned, but the remainder of the building is not. Duplicate computer tapes, updated weekly, are stored in a well-protected vault in another part of the city. A monthly fee is paid for the tape storage.

ABC owns several small trucks that are used to move lumber about its own premises and for local pickup and delivery. The trucks are kept in the lumber storage shed when they are not in use. Several lift trucks are used for moving lumber and other heavy items in and around the factory and lumber storage area.

A freight elevator moves materials, finished furniture, lift trucks, and other equipment between floors in the factory. Automatic grillwork gates have been installed to prevent workers from falling into the elevator shaft. There is no passenger elevator. Workers use either the freight elevator or the open stairwells when moving between floors. Steel fire escapes have been installed on the exterior of the building to facilitate evacuation of the upper floors in case of fire or other catastrophe.

ABC Manufacturing Company is wholly owned by its president, Mr. Carpenter, who is the grandson of the company's founder. Mr. Carpenter is considered wealthy by local standards. However, his fortune consists almost entirely of the stock of ABC Manufacturing Company, and he is dependent on his salary and company dividends for his livelihood. Nearly all of ABC's operating profit has been paid in dividends in recent years. Consequently, the company has only a modest cushion of liquid assets in excess of its operating needs. Profits have been consistent, but they have been relatively low because of the inefficiency of the present factory facilities. Although the dollar amount of ABC's sales has increased steadily because of price increases, the physical volume of sales has remained almost constant over the past several years. The lack of growth has resulted primarily from two factors. First, the present plant cannot accommodate greater production because of space limitations, and there is no available land adjacent to the plant to permit expansion. Second, the long training period required for new employees prevents rapid expansion of production. Mr. Carpenter has considered building a new factory at another location in the same town. However, such a move is not financially feasible unless the present plant can be sold, and no prospective purchasers have been

Table 14-1
ABC Company Building Values

| Building | Actual Cash Value | Replacement Cost |
|---|---|---|
| Factory | $1,500,000 | $2,500,000 |
| Lumber shed | 150,000 | 200,000 |
| Power plant | 230,000 | 300,000 |
| Kiln | 270,000 | 320,000 |

found. A move to any location outside its present hometown would not be practical because of the company's dependence on its well-trained craftsmen. The sales manager estimates that both the dealer network and sales could be increased by 30 percent over the next five years if production facilities could be provided.

The cost to rebuild the present factory building in its present form would be $2.5 million. However, Mr. Carpenter has indicated that he would not build a new mill-type building to replace the current structure because of the high cost of the thick brick walls and heavy timber interior construction. A new, one-story, noncombustible masonry and steel building of comparable floor area would cost approximately $2 million to build and would be more efficient for ABC's purpose. The actual cash value of the present factory building is estimated to be $1.5 million, and ABC has insured it for that amount against fire and the extended coverage perils. The lumber storage shed, power plant, table and lumber kiln are also insured for their actual cash values. Table 14-1 shows the actual cash value and replacement cost for each of the structures.

ABC's annual premium for fire and extended coverage insurance is $136,000, including the coverage for contents of the buildings and for lumber stored in the yard. There have been no fire or extended coverage losses during the past five years and only minor losses prior to that time. The company's estimated annual workers' compensation premium is $165,000. Loss experience has been fairly consistent from year to year. Based on past experience, with adjustments for inflation and current benefit levels, normal losses of $104,000 can be expected.

The most frequent claims have been small and have resulted from such minor injuries as splinters in hands, several minor dermatitis cases, and sawdust or metal particles in the eyes. The more serious injuries have included back strains and loss of fingers in power saws and other power tools. ABC now has a 15 percent debit under the workers'

compensation experience rating plan. (That is, they are paying a rate 15 percent greater than manual rates.)

17

Case 2



# BITE-O-BURGER COMPANY

## Description of Operations

The Bite-O-Burger Company is a publicly held corporation. It owns and operates 843 fast-food restaurants located in eleven states. The restaurants feature a limited menu consisting of hamburgers, french fried potatoes, fried chicken, chili, related food items, and nonalcoholic beverages. The restaurants vary in size, but each is located in a free-standing building and surrounded by customer parking areas. All of the buildings were built to Bite-O-Burger's plans and specifications and share enough architectural characteristics to make them easily recognized as units of the chain. All have forced air heat and are air conditioned.

At current prices, the average replacement cost of the restaurants is estimated at $125,000 per unit for the building and $100,000 for the equipment. Because of differences in size, the replacement cost, including building and equipment, ranges from $175,000 for the smallest

restaurant to $300,000 for the largest. The average actual cash value is $205,000 for building and equipment combined. The restaurants vary in age from a few days to approximately twenty years. All of them are owned by Bite-O-Burger, but the newer ones are subject to substantial mortgages.

The home office of the company is located in leased space in a building in the business district of a midwestern city. Bite-O-Burger occupies the upper three floors of the thirty-story building. The company's data processing center is located on the top floor. All of the computer equipment is leased from the manufacturer.

Bite-O-Burger also occupies a leased warehouse near the home office. It is used for storage and distribution of supplies (paper cups, wrapping materials, etc.) and nonperishable food items. Perishable food items are purchased from local suppliers near the restaurants in which they will be used, and they are delivered directly to the restaurants by the suppliers. There is no refrigeration equipment at the warehouse, but each restaurant has a large, walk-in refrigerator. Items from the company warehouse are distributed to the individual restaurants by a fleet of thirty owned tractor-trailer units. The same units also transport goods from the suppliers to the central warehouse when truck-load quantities are purchased. Smaller lots are shipped by common carrier FOB point of shipment.

The values of the contents at the warehouse and home office are $15,750,000 and $3,200,000, respectively. Values at both locations are relatively constant throughout the year.

Bite-O-Burger advertises extensively in newspapers in the cities in which it has several restaurants. Many of its advertisements feature endorsements of its products by prominent athletes and theater personalities. Some advertisements feature pictures of local people and their favorable comments on the company's food and service. The company also sponsors softball and bowling teams in some cities as a part of its public relations program.

Bite-O-Burger's profit and loss statement and an abbreviated balance sheet for last year are shown in Figures 14-1 and 14-2. The company's operating results for last year were typical of past years, but sales and assets have been growing at a rate of approximately 20 percent per year.

Bite-O-Burger's fire and extended coverage losses for the past five years are shown in Table 14-2.

The quotations shown in Table 14-3 have been obtained for fire and extended coverage protection—blanket on buildings and contents on an actual cash value basis. Bite-O-Burger is well aware of the fire exposures associated with restaurants. Extensive fire control equipment

Figure 14-1
Bite-O-Burger Company Profit and Loss Statement

Profit and Loss Statement

| | | | |
|---|---|---|---|
| Sales | | | $345,630,000 |
| Cost of materials and supplies | | $102,049,359 | |
| Wages and salaries | | | |
| Officers and key employees | $ 13,306,755 | | |
| Other employees | 107,663,740 | 120,970,495 | |
| Utilities | | 8,641,258 | |
| Maintenance and repairs | | 9,506,873 | |
| Taxes (other than income taxes) | | 6,920,101 | |
| Insurance | | 4,320,376 | |
| Advertising | | 14,368,897 | |
| Depreciation | | 10,176,438 | |
| Miscellaneous | | 3,974,744 | 280,928,541 |
| Net profit before taxes | | | $ 64,701,459 |
| Income taxes | | | 38,458,674 |
| Net profit after taxes | | | $ 26,242,785 |

has been installed in the kitchens of all units, and especially in the range hoods and in the cooking areas.

Figure 14-2
Bite-O-Burger Company Balance Sheet

Balance Sheet

| | | | | | |
|---|---|---|---|---|---|
| **Assets** | | | **Liabilities** | | |
| Cash | $12,357,821 | | Accounts payable | $20,269,847 | |
| Inventory | 10,879,643 | | Short-term bank loans | 21,764,169 | |
| Buildings† | 73,730,466 | | Salaries and wages payable | 2,419,410 | |
| Land | 39,269,469 | | Mortgages on real estate | 66,327,883 | |
| Furniture, fixtures and equipment† | 54,757,469 | | Miscellaneous liabilities | 831,779 | |
| Motor vehicles† | 896,010 | | Total liabilities | | $111,613,088 |
| Miscellaneous assets | 412,210 | | | | |
| Total assets | | $192,303,088 | **Capital and Surplus** | | |
| | | | Common stock | $20,000,000 | |
| | | | 8% preferred stock | 30,000,000 | |
| | | | Paid-in surplus | 10,000,000 | |
| | | | Earned surplus | 20,690,000 | |
| | | | Total capital and surplus | | $ 80,690,000 |
| | | | Total liabilities, capital, and surplus | | $192,303,088 |

†After deduction for depreciation.

Table 14-2

Bite-O-Burger Company Fire and Extended Coverage Loss History

| Year | Number of Losses | Total Amount | Largest Loss |
|---|---|---|---|
| 1 (most recent) | 20 | $ 17,860 | $ 8,843 |
| 2 | 15 | 11,760 | 6,981 |
| 3 | 17 | 256,549 | 247,317 |
| 4 | 12 | 9,336 | 2,178 |
| 5 | 11 | 18,393 | 11,423 |
| | | $313,898 | $276,742 |

Table 14-3

Bite-O-Burger Company Fire and Extended Coverage Insurance Quotations

| Deductible Amount | Annual Premium |
|---|---|
| $100 per occurrence | $769,000 |
| $5,000 per occurrence | 630,500 |
| $10,000 per occurrence | 545,900 |
| $25,000 per occurrence | 439,330 |
| $100,000 per occurrence | 357,600 |
| $100,000 annual aggregate | 461,000 |
| $250,000 annual aggregate | 243,000 |

18

Case 3



# PANACEA PHARMACEUTICAL COMPANY

## Description of Operations

The Panacea Pharmaceutical Company is a publicly owned corporation. Its common stock is traded on a major stock exchange. The company is engaged in the manufacture and distribution of both ethical drugs (sold only by prescription) and patent medicines sold over the counter by drug stores and other outlets. Some of the ethical pharmaceuticals contain narcotics, and substantial amounts of narcotics are stored in the manufacturing plant for use in the preparation of such products.

Although Panacea makes and sells over 100 different drug products, approximately 30 percent of sales and a slightly higher percentage of profits are derived from one product, an ulcer remedy. There are other similar ulcer remedies on the market, and management concedes that some of them are equally effective. However, Panacea's product was the first on the market and has acquired a large following among doctors, who usually prescribe it by Panacea's brand name.

Management is concerned that any prolonged absence of their ulcer medicine from the market would cause doctors to prescribe another brand and they might not return to Panacea's brand when it becomes available again. The manufacturing process for the product is relatively simple. Only stock machinery is needed. Replacement machinery would be available immediately, and several buildings suitable for the purpose are also readily available. The manufacture of the product (and all other company products) could be subcontracted to other pharmaceutical manufacturers who specialize in making such products on a contract basis. However, about thirty days would be required to negotiate contracts, obtain packaging materials, and make other necessary arrangements to farm out production.

The formula for the ulcer medicine includes one chemical compound that is manufactured in only one chemical plant in the world. The reason

for its limited production is lack of demand. Panacea uses about one-half of the amount produced. The compound is not difficult to produce and is not protected by patents. Panacea's production manager estimates that Panacea could install the necessary equipment, obtain necessary materials, and begin producing the compound themselves in about sixty days. However, they would prefer to continue purchasing it from the present supplier if possible. All other materials used in Panacea products are readily available from several sources.

Panacea's sales are growing at a rate of about 20 percent per year, mostly as the result of the introduction of new products. The company has extensive research facilities and has developed and introduced an average of five new products per year over the last six years. Many other potential products are developed and tested but are not introduced because they prove to be ineffective, dangerous, or both. The testing process for some products is long and complex. Some products may involve several years of testing on dogs, primates, or other relatively long-lived animals, possibly followed by testing on human volunteers. Voluminous records are accumulated during such tests. The records must be retained for many years for use in licensing applications and defense of products liability claims, for use in future research projects, and for other purposes.

The research records are kept in fire-resistive filing cabinets in the records room of Panacea's research center. The center also houses research laboratories, offices for research personnel, and cages for experimental animals. It is located in a sprinklered, fire-resistive building adjacent to Panacea's factory.

The factory building also is fire resistive and is sprinklered in all areas except the clean room. The clean room is used for manufacturing and packaging processes that require complete sterility. It has its own air conditioning system with special filtering equipment to eliminate dust or other potential contaminants. Other equipment is also provided to maintain the sterile atmosphere. Even very slight contamination of the clean room would require that production be discontinued for several days until sterility could be reestablished. All workers in the clean room must wear special sterile uniforms and surgical masks.

Both the factory and the research center are five years old. They are surrounded by a carefully maintained lawn, and are separated from the nearest building by a distance of 200 feet. They are located in a medium-sized city with excellent public fire protection and water supply.

Panacea's products are marketed through company branches. Each branch consists of an office and a warehouse. There are twenty-three branches in the United States and eighteen branches in Canada, Europe, and Latin America. Each branch has several sales representatives who call on doctors, hospitals, pharmacists, and other retailers of drug

products. The sales representatives at the foreign locations are citizens of the countries in which the branches are located. However, most of the branch managers, assistant branch managers, and branch sales managers are U.S. citizens. The company's products are offered for sale in approximately 40,000 retail establishments here and abroad.

The branches are located in leased quarters, but the office and warehouse equipment and drug inventory total about $350,000 at each location. Most shipments to domestic branch locations, as well as incoming shipments of materials, are made by common carrier trucks and railroads. Foreign shipments usually go by ship. However, air freight shipments, both domestic and foreign, are sometimes made if great speed is necessary. The only motor vehicles owned by Panacea are twelve private passenger cars used by executives and messengers and three light trucks used for local pickup and delivery. All of them are at the home office. Branch office officials and sales representatives use their personal automobiles for business-related travel. Panacea does not own any boats or aircraft, but small airplanes, with crew, are sometimes chartered for executive transportation. The sales vice president, a licensed pilot, sometimes uses his personal airplane on company business.

Panacea's products are advertised extensively. Ethical pharmaceuticals are advertised primarily by direct mail to doctors and in medical journals. These advertisements tend to be technical and medical in nature. Advertisements for patent medicines appear on radio and television and in consumer magazines and newspapers. Such advertisements are nontechnical and are designed to appeal to consumers. Endorsements by prominent persons are sometimes used.

Panacea is in an exceptionally strong financial condition. Profits have been very satisfactory and have been quite stable over the past ten years. The company is accumulating a fund from retained earnings to expand its production facilities. The fund, which now stands at $10 million, is invested in liquid securities. The securities that are nonnegotiable are kept in a safe in the office section of the factory building. Management estimates that present production facilities are adequate for the next five years and that the fund will be adequate by the end of that period to pay for the needed expansion. Panacea does not have any long-term debt, and its only short-term debt consists of accounts payable to suppliers.

The present manufacturing vice president who has occupied that position for five years has been more aware than his predecessor of the importance of loss control. The safety and industrial hygiene programs he has initiated have been very successful in reducing employee injuries. The company's workers' compensation premiums, losses, and experience rating modifications for the last five years are shown in Table 15-1.

The experience rating modification for the current year is a credit

**Table 15-1**
Panacea Pharmaceutical Workers' Compensation History

| Year | Premium | Losses Number | Losses Amount | Experience Debit (+) or Credit (−) |
|---|---|---|---|---|
| 1 | $110,000 | 57 | $87,000 | + 30% |
| 2 | 120,000 | 48 | 73,000 | + 23% |
| 3 | 125,000 | 46 | 75,000 | + 18% |
| 4 | 125,000 | 47 | 71,000 | + 12% |
| 5† | 127,000 | 43 | 66,000 | + 3% |

† Most recent completed year.

**Table 15-2**
Panacea Pharmaceutical Estimated Cost of Workers' Compensation Retention Program

| | |
|---|---|
| Losses | $52,000 |
| Excess insurance | 9,200 |
| Self-insurer bond | 1,900 |
| Taxes | 600 |
| Service fee | 10,200 |
| Total | $73,900 |

of 2 percent. The increasing premium reflects increases in manual rates and payrolls, which have more than offset the improved experience modification. The factory, research center, and home office payrolls account for about 75 percent of the workers' compensation premium. Panacea has been considering the feasibility of retaining the workers' compensation exposure for all employees in its home state. Commercial insurance would be continued for all other employees because of the expense and inconvenience of qualifying as a "self-insurer" in all states. A risk management consultant has prepared an estimate of the annual cost of a retention program for Panacea's home state employees, as shown in Table 15-2.

The estimated annual workers' compensation premium for the same part of Panacea's payrolls is $96,000. The excess insurance premium would provide coverage for $2 million in excess of aggregate losses of $125,000 in one year. The self-insurer bond is required by the state to

guarantee the payment of benefits, and the taxes are for the support of the state workers' compensation board. The service fee would be paid to a service company for loss adjustment, loss control, and administrative services furnished by that firm. The service company's staff is capable of providing those services on the same level now provided by the insurance company. Panacea does not have any employees qualified to provide claims service. It has only one safety engineer on its staff